TAMPERE UNIVERSITY OF TECHNOLOGY

MATTI MATTILA

MICROPHONE ARRAY VERIFICATION FOR BROADBAND SODAR

Master of Science Thesis

Examiners: Professor Timo Hämäläinen
Professor Ari Visa
Examiners and topic approved by the Faculty Council of the Faculty of Computing and Electrical Engineering on 5.2.2014

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO
Signaalinkäsittelyn ja tietoliikennetekniikan koulutusohjelma
MATTI MATTILA: Microphone Array Verification for Broadband Sodar
Diplomityö, 65 sivua
Toukokuu 2014
Pääaine: Sulautetut järjestelmät
Tarkastajat: Professori Timo Hämäläinen ja professori Ari Visa
Avainsanat: tutka, äänitutka, sensoriryhmä, keilanmuodostus, kalibrointi

Monet seurantajärjestelmät kuten tutka (engl. radar) toimivat lähettämällä signaalia ja kuuntelemalla kaikuja tästä lähetetystä signaalista. Tutkan tapauksessa lähetetty signaali on radioaalto. Ääniaaltoja voidaan myös käyttää kuten radioaaltoja tutkassa ja väliaineen ollessa ilma kutsutaan laitetta äänitutkaksi (engl. sodar). Tutkassa vastaanotto ja lähetys tapahtuvat antenneilla, kun taas äänitutkassa mikrofoneilla ja kaiuttimilla. Kohteen etäisyys saadaan laskemalla aika signaalin lähetyksen ja kaiun vastaanoton välillä. Kohteen suunta taas saadaan selvitettyä kääntämällä vastaanottavia sensoreita.

Tämä diplomityö on osa isompaa kokonaisuutta, joka tähtää halpaan tutkan toiminnallisuuden testausjärjestelmään. Työn tavoitteena oli selvittää riittääkö halpojen mikrofonien tarkkuus laajakaistaisen äänitutkan toteuttamiseen. Tämän selvittämistä varten tutustuttiin ensimmäisenä äänitutkan taustalla tapahtuviin ilmiöihin. Koska äänitutkaan liittyvää kirjallisuutta on tehty verrattain vähän esiteltiin se tutkan teorioiden avulla. Työssä käsitellään myös keilanmuodostusta. Keilanmuodostus on signaalinkäsittelymenetelmä sensoriryhmän spatiaalisen toiminallisuuden hallintaan. Keilanmuodostus parantaa suorituskykyä, mahdollistaa digitaalisen keilankääntämisen sekä laajan taajuuskaistan hallinan.

Itse mikrofonien varmentaminen tapahtui tekemällä teoreettiset simulaatiot sekä käytännön mittaukset ja vertailemalla näitä keskenään. Ensimmäisten mittausten jälkeen tuli jo selväksi että ostetut mikrofonit riittävät äänitutkan toteuttamiseen. Tällaisenaan mikrofonien epäideaalisuudet näkyvät ryhmän keilakuviossa. Tämän takia mikrofonit pitää kalibroida ennen käyttöä ja siksi tämä työ esittelee myös mikrofonikalibrointimenetelmän, jonka jälkeen keilakuvio on verrattain lähellä ideaalitapausta. Työssä mitattiin ja simuloitiin myös digitaalista keilankääntöä sekä laajakaistaista keilanmuodostusta. Molemmissa tapauksissa mittaustulokset ovat lähellä ideaalisia simulaatioita. Täten työ todistaa, että ostetut mikrofonit ovat riittävät laajakaistaiseen äänitutkasovellukseen.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Signal Processing and Communications Engineering
MATTI MATTILA : Microphone Array Verification for Broadband Sodar
Master of Science Thesis, 65 pages
May 2014
Major: Embedded Systems
Examiners: Professor Timo Hämäläinen and Professor Ari Visa
Keywords: radar, sodar, sensor array, beamforming, calibration

Many surveillance systems, such as radar (radio detection and ranging), work in a way that a signal is transmitted and its echoes are detected. In the case of radar, the transmitted wave is a radio wave. When sound waves are used instead of radio waves and the medium is air, the system is called sodar (sonic detection and ranging). In radar, antennas are the receiving and transmitting sensors, while in sodar the sensors are microphones and speakers. The range of the target can be determined by calculating the time between the transmission of the signal and the reception of the echo. The bearing of the target can be determined by rotating the receiving sensors.

This thesis is a part of a project that aimed to develop a cheap radar-emulation environment. The main objective of the thesis was to verify if inexpensive microphones are feasible for a microphone array implementation for the broadband sodar. First, the phenomena behind sodar were studied. Since literature on the topic is scarce, sodar is explained with the theories of radar. The concept of the beamforming was also studied. Beamforming is a signal processing method for controlling the spatial properties of a sensor array. Beamforming enhances the performance, and enables both digital beam steering and the control of the broad frequency band signals.

The microphone verification was accomplished with theoretical simulations and practical measurements, and comparing these. After the first measurement it already became clear that the microphones studied are capable for the sodar implementation. However, the microphone unidealities are apparent in the beam pattern of the measured array. For this reason, the microphones should be calibrated before use. Thus, this thesis introduces a microphone calibration method after which the beam pattern becomes nearly ideal. The digital beam streering and broadband beamforming were also simulated and measured. In both cases, the measurement result is close to the ideal models. Thus, this thesis proves that inexpensive microphones are capable for broadband sodar implementation.

PREFACE

This Master of Science Thesis was done for the Multimedia and Data Mining research group of the Department of Signal Processing at Tampere University of Technology. Luckily for me, my summer job application for the Department of Signal Processing stood out and I was given an opportunity to join the research group for the purpose of writing this thesis.

There are a great many people I have had the privilege to meet during my studies and who deserve my thanks. I would like to thank fellow workers of MMDM group for their help and support. I would especially like to thank Juha Jylhä and Tomi Nihtilä. Tomi has been helping me with practically everything and been willing to explain matters whenever I have been struggling with understanding them. Juha has instructed my work from the beginning, and I am really grateful for his understandable explanations of difficult concepts. One thing he said was especially encouraging: "This is not an easy topic so don't worry too much if it doesn't open up on the first try". My thanks go also to Audio Research Team of Department of Signal Processing for borrowing equipment to our measurements, and to Pasi Pertilä of the same team for several practical tips for my work. I also thank Professor Timo Hämäläinen for volunteering as an examiner for my thesis and Professor Ari Visa as being an examiner on behalf MMDM group.

Finally, I would like to thank my family. Their love and support have been invaluable. Even the never-ending disbelief of my parents about my graduation and finding a real job has managed to push me forward. Last but not least, I would especially like to thank Suvi who proofread my thesis and whose endless support has helped me through the workload of the last few years.

Tampere, May 15th, 2014

Matti Mattila

CONTENTS

LIST OF ABBREVIATIONS

FIR Filter	Finite Impulse Response Filter
SODAR	Sonic Detection And Ranging
RADAR	Radio Detection And Ranging
SONAR	Sound Navigation And Ranging
RTT	Round Trip Time
CW	Continuous Wave
FM	Frequency Modulation
TDOA	Time Difference of Arrival
RF	Radio Frequency
LO	Local Oscillator
IF	Intermediate Frequency
SNR	Signal to Noise Ratio
HPBW	Half-Power Beamwidth
FNBW	First Null Beamwidth
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
ULA	Uniform Linear Array
DS	Delay-and-Sum
FS	Filter-and-Sum
CDB	Constant Directivity Beamforming
DC	Dolph-Chebyshev
ASIO	Audio Stream Input/Output
FFT	Fast Fourier Transform

1. INTRODUCTION

Propagating waves have been used for target detection for over a hundred years. The most known devices are radar and sonar: radio detection and ranging which uses radio waves for detection, and sound navigation and ranging which uses sound waves for underwater surveillance. Sound waves can also be used in the air: the device is called sonic detection and ranging (sodar). All work in a similar way; a signal is transmitted and echoes from the possible targets are listened. The range is determined by measuring the time between the transmission of the signal and reception of the echo from a target. Unfortunately, in most of the cases, the range alone is not enough. In addition to the distance of the target, the bearing of the target is an important piece of information. An antenna radiates most of its power towards some specific direction, which is called an antenna directivity. The conventional antenna has a fixed directivity pattern. Thus, the position or direction of the target is traditionally determined by physically turning the receiver antenna towards a direction of interest. Directivity can be improved by using multiple antennas to form an antenna array. In addition to improved directivity, this enables a digital steering towards the desired direction. [1; 2]

The beam pattern shape is heavily frequency dependent. If the frequency of the transmitted wave is low, then the main lobe will be wide. Conversely, when wave's frequency is high, the width of the main lobe becomes narrower. This is a physical property of antennas, and for this reason there are limited means to control this frequency dependency. However, certain beamforming methods are suitable for frequency dependency control. [3]

Radar has the widest range of applications. It is used, for example, in different vehicles, military applications and measurement applications. [4] Sonar is broadly used in several waterborne vessels. Its uses include avoiding obstacles, such as rocks, and finding desired targets, for example fish. [5] Sodar can theoretically be used for the same applications as radar, but because of the nature of sound waves, some limitations exist. For example, air affects sound waves much more than radio waves. Thus, sodar is only usable for short distances. Sound waves also scatter back from air turbulences. Because of this, the most typical application for sodar is, for example, creating a wind profile for wind power turbines. [6]

1.1 Project Background

This Master of Science thesis was done for the Multimedia and Data Mining research group of the Department of Signal Processing at Tampere University of Technology. The group has a solid know-how on radar theories and has done a great deal of research on the field. For the MMDM, my work was a part of a project to develop an inexpensive radar-emulation environment. At a certain point, the idea of using sound waves for existing radar methods was born. The reason to use sound waves was partly due to the interest to see how well sound waves could be used with radar techniques. Radar systems are both expensive and complicated, and, furthermore, the use of frequency bands is regulated. Thus, if sound waves perform well with the radar techniques, an acoustic radar system could be used for testing the new methods before the verification with radar.

As was mentioned previously in this chapter, the directivity of a single antenna is fixed after its production. However, the directivity can be enhanced by using multiple antennas for detecting the same signal. Expensive components would ensure good directivity, but if less expensive materials are used, the only practical solution is to form an array. Thus, this thesis offers a background for using multiple microphones to enhance the performance of directivity for a sodar system.

Tomi Nihtilä had a major role in the sodar project. He was responsible for designing and implementing the research system for the project. This thesis also documents the essential parts of the implemented research system.

1.2 Objectives

The main objective of the thesis is to verify if the inexpensive consumer microphones could be used for a sodar project. It also aims to offer insights to the theory behind microphone arrays. Because this thesis is part of a project for radar-emulation environment, the radar and sodar theory are discussed in more depth than would be necessary for the measurement related to the microphone arrays. Theoretically microphone and antenna arrays behave in similar ways, and my hope is that the background work done for the thesis can be used in different remote sensing studies as well.

In the sodar system, microphones work as receivers and speakers as transmitters. First, it was necessary to make sure that the ordered microphones can be used for building this kind of system. As the directivity of the array varies with frequency, the thesis also aims to find ways to control frequency dependency. This was done in MATLAB environment. The first step was a theoretical simulation. Then, the same operation was done by measuring the performance with the actual microphone array. It was expected that the performances between microphones are not exactly

the same. Thus, one of the main goals of this thesis was to study if some calibration method could be used to correct these microphone unidealities.

1.3 Thesis Structure

The thesis is structured in the following way. Chapter 2 concentrates on theory of both radar and sodar. The chapter covers the essential theory on the waveforms and detection. In Chapter 3 the background of the array theory is explained. The chapter begins with fundamental cases and then proceeds towards more complex ones, such as the case with multiple frequencies. A few useful signal processing techniques and different array geometries are also introduced. Chapter 4 describes the used research system designed and built by Tomi Nihtilä, because it has previously not been documented. It documents specification for the equipment used, and explains the operation of the implemented sodar. It also introduces a calibration method for a microphone array. Chapter 5 presents the results of the simulations and measurements conducted for this thesis by the author. In conclusion, Chapter 6 sums up the key ideas of the thesis, offers some discussion of the results and briefly describes ideas for future development.

2. THEORY ON SOUND, DETECTION AND RANGING

This chapter discusses sodar and the main theory behind it. In this thesis, sodar is defined in analogy with radar. This means that sodar can be seen as "an acoustic radar"; in other words, an application which uses sound waves for detection. It is similar to sonar, but sodar is considered to use air as the medium, whereas sonar uses water. [7] For the purpose of understanding sodar, some basic theory of sound will be covered in the following sections. In addition, because radar better presented in the literature, sodar is explained through the theory of radar.

2.1 Radio Detection and Ranging

Radio detection and ranging, or radar, is a widely used surveillance device. Radar's usage breakthrough happened during Second World War. After the war, it has become an every day equipment in vehicles. It is used in measurements, and has also military uses. This section will discuss fundamental radar theory.

2.1.1 Radar Basics

Radar is a device which uses electromagnetic waves for detecting targets. Because most objects reflect electromagnetic waves, radar can transmit a wave and then try to detect an echo from the target. A typical block diagram of pulsed radar is depicted in Figure 2.1. The transmitter and receiver can be separate from each other. Or, more commonly, they may share a common antenna. Using the same antenna for both transmitting and receiving causes some interference. To avoid this problem, the signal is often transmitted in pulses. In both cases, the main idea remains the same: an electromagnetic signal is transmitted, an echo is received and the target information is either further processed, or, more simply, merely shown on the display. While the echo is not interesting on its own, it nonetheless enables extraction of the target's information. Typically, the most important piece of information is the location which includes a range and direction. The range can be determined by measuring the time it takes for the signal to reach the target and return. Radio waves travel at the speed of light. Thus, the range can be calculated as

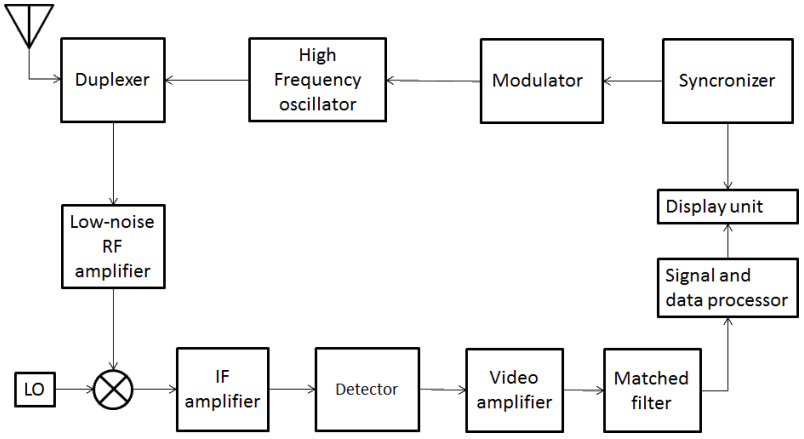

Figure 2.1: Radar block diagram. [8]

$$R = \frac{1}{2} \times RTT \times c \tag{2.1}$$

in which RTT is the round trip time and c is the speed of light. In pulsed radar, the range is achieved by transmitting a pulse and calculating the time to the reception of the echo. However, in continuous wave radars, for example, in which both transmission and detection are active simultaneously, this is not an effective technique. These CW radars often measure range with *frequency-modulation* ranging. This means that frequency of the transmitted wave is varied, and the range can be determined by observing the time lag between the modulation of the transmitted signal and the corresponding modulation of the echoes received. The idea is illustrated in Figure 2.2. The range in FM ranging can be calculated as

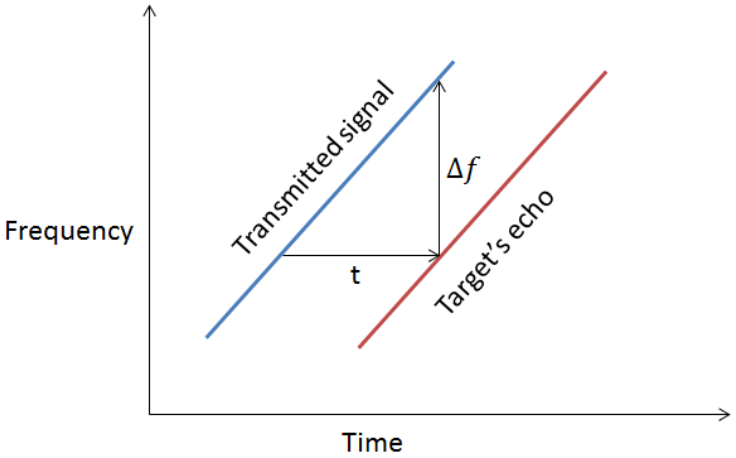

Figure 2.2: Illustration of FM ranging.

$$R = \frac{1}{2} \times c \times t \tag{2.2}$$

and t can be written as

$$R = \frac{1}{k} \times \Delta f \tag{2.3}$$

in which $k = \frac{\Delta f}{\Delta t}$ is the rate of change. [9]

FM can also be used in pulsed radars to improve the range resolution. This method is often called *pulse compression*. Range resolution of pulsed radar can be written as

$$\delta_r = \frac{c}{2B} \tag{2.4}$$

in which B is the bandwidth of the transmitted waveform. For simple pulsed radar system, $B = \frac{1}{T}$ in which T is the transmitted pulse duration, and range resolution can then be written as

$$\delta_r = \frac{cT}{2}. \tag{2.5}$$

In pulse compression system, the transmitted waveform is modulated so that $B \gg \frac{1}{T}$. By writing $\tau = \frac{1}{B}$, the resolution can be written as

$$\delta_r = \frac{c\tau}{2} \tag{2.6}$$

in which τ is the effective pulse duration of the system after pulse compression. This means that the system can use pulse duration T, while having the range resolution equivalent system with the pulse duration τ, where $T \gg \tau$. In other words, this allows the system to use a long pulse with low peak power, and attain the resolution of a short pulse which would normally need high peak power. [10]

The echo also reveals the direction of the target. The direction is often measured in terms of the angle between the line of sight to the target and a horizontal reference direction. Angle is usually split into its horizontal and vertical component, also called *azimuth* and *elevation*. DOA estimation can be done by turning antennas towards the direction of interest. Antennas can be turned physically (this is the traditional method). Alternatively, the antenna array can be formed, which enables digital steering of the detecting beam. Forming an antenna array is covered in Chapter 3. [9]

2.1.2 Radar Block Diagram

The radar block diagram in Figure 2.1 can be divided into transmission and reception. In this case, the duplexer and the antenna are used for both transmission and reception. The transmission consists of *synchronizer*, *modulator*, *high frequency oscillator*, *duplexer* and *antenna*. The synchronizer is also called a trigger circuit. It generates pulses to the modulator at a regular repetition rate which is used to generate high frequency pulses. The synchronizer provides rectangular pulse and acts as a supply voltage to high frequency oscillator. The high frequency oscillator is switched on for very short intervals to generate high frequency pulses. These pulses are transmitted at regular intervals by the synchronizer. The oscillator produces high frequency pulses with very high power. If the same antenna is used for transmitting and receiving, a duplexer is needed to choose which of the operations is in progress. [8]

Reception consists of *duplexer*, *low-noise RF amplifier*, *local oscillator*, *mixer*, *IF amplifier*, *detector*, *video amplifier*, *matched filter*, *signal and data processor* and *display unit*. A low-noise RF amplifier is used to amplify the signal to a desired level in which RF stands for *radio frequency*. This amplified signal is then mixed with a signal generated by local oscillator (LO). This converts the RF signal into the lower frequency level called the *intermediate frequency* (IF). Then, the IF signal is amplified by an IF amplifier. The amplified IF signal is demodulated by detector. Then the signal is amplified by a video amplifier, which strengthens the signal with large bandwidth to preserve the shape of the received pulse. The signal is then filtered with a matched filter, which is a filter that optimises the signal to noise ratio (SNR). The effect of the medium to a propagating signal cannot be perfectly predicted: thus, a true matched filter cannot be implemented. However, a practical filter can be used to filter noise while retaining the signal. The signal and data processor is used to process the signal data. The processing can be, for example, target echo processing, interference suppression and incrementing SNR. Detection information for deciding whether the target is present or not is also provided. Information of the range and the possible Doppler shift also acquired as well. Finally, the result is shown to the operator on the display. [8]

2.1.3 Radar Frequencies and Radio Wave Propagation

Radar can be used within a broad band of frequencies. Most commonly it is used from 1 GHz to 40 GHz, but it can be used up to 3 PHz (10^{15}), which is already in the frequency band of visible and ultraviolet light. The official way is to designate letters to different frequency bands. The frequency spectra are regulated by governments. IEEE standard frequency bands are presented in Table 2.1.

Table 2.1: IEEE standard frequency bands

HF = 3 ... 30 MHz
VHF = 30 ... 300 MHz
UHF = 300 ... 1000 MHz
L = 1 .. 2 GHz
S = 2 ... 4 GHz
C = 4 ... 8 GHz
X = 8 .. 12 GHz
K_u = 12 ... 18 GHz
K = 18 ... 27 GHz
K_a = 27 ... 40 GHz
V = 40 ... 75 GHz
mm = 110 ... 300 GHz

The criteria for band selection depends on the use case. If the radar needs to detect objects which are far away and discrimination is not an issue, lower frequency should be used. Alternatively, adequate discrimination needs high frequency and, thus, detection range remains short. Weather conditions also become an issue with higher frequencies. [11]

The propagation of radio waves depends on the medium. Loss is reasonably low in air or free space, and higher in water or inside earth. All media are characterised with *permittivity* ϵ, *permeability* μ and *conductivity* σ. Permittivity describes how the medium affects the electric field, and permeability describes effect on the magnetic field. Conductivity describes medium's ability to conduct electricity. The radio wave velocity is related to the preciding parameters, and can be expressed as:

$$V = \frac{1}{\sqrt{\epsilon\mu}} = \frac{1}{\sqrt{\epsilon_r\mu_r \times \epsilon_0 \times \mu_0}} \tag{2.7}$$

in which V is the velocity of radio wave, ϵ and μ are media permittivity and permeability, ϵ_r and μ_r are free space permittivity and permeability, ϵ_0 and μ_0 relative permittivity and permeability constants. There are several medium-related phenomena which affect radio wave propagation, such as reflections and multipath radio links, atmospheric refraction, curvature of radio path and K-factor, earth based and elevated radio ducts, diffraction and obstruction loss, free space loss, atmospheric attenuation due to absorptions, the scattering of radio waves, the depolarisation of radio waves, sun spot effects and magnetic storms. [12]

2.1.4 Antenna Basics

A radio antenna is defined as the structure associated with the region of transition between a guided wave and a free-space wave, or vice versa. In other words, the antenna is a transducer between these two wave types. Antennas radiate mostly towards some specific direction: this phenomenon is called an antenna *directivity*. Radiation pattern (also called a beam pattern) is an illustration of this directivity. An illustration of an antenna radiation pattern is depicted in Figure 2.3. Three important phenomena are marked on the figure. The *main lobe* represents the sector with the highest radiation towards the intended direction. A practical antenna cannot radiate its power only to a single direction: because of this, *side lobes* are formed. An antenna also radiates to its back; these lobes are called *back lobes*. Between the lobes, there are spaces where the field reaches zero; these spaces are called *nulls*. Several factors are beneficial to antenna system design. A few simple but useful ones are half-power beamwidth (HPBW), beamwidth between first nulls (FNBW), directivity (D) or gain (G). HPBW means the angular beamwidth at the half-power level, while FNBW is the angular distance to first nulls of the pattern. The directivity of an antenna is equal to the ratio of the maximum power density $P(\theta, \phi)_{max}$ to its average value over a sphere, as observed in the far field of an antenna. This can be written as

$$D = \frac{P(\theta, \phi)_{max}}{P(\theta, \phi)_{av}}. \tag{2.8}$$

It is a dimensionless ratio ≥ 1. Average power density can be written as

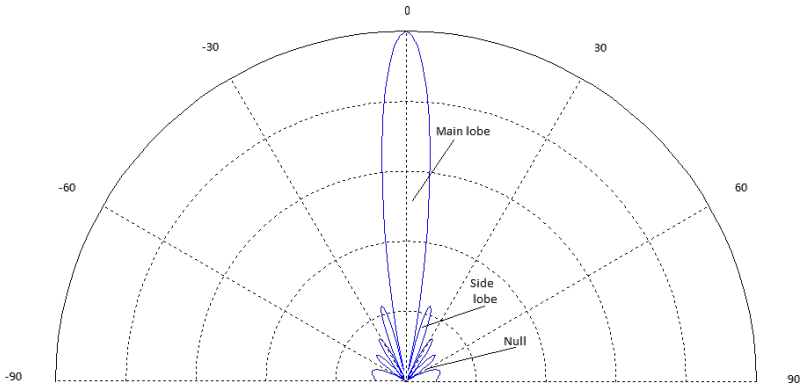

Figure 2.3: Radiation pattern of a directive antenna in polar coordinate system.

$$\begin{aligned} P(\theta,\phi)_{av} &= \frac{1}{4\pi}\int_{\phi=0}^{\phi=2\pi}\int_{\theta=0}^{\theta=\pi} P(\theta,\phi)\sin\theta\ d\theta\ d\phi \\ &= \frac{1}{4\pi}\iint_{4\pi} P(\theta,\phi)d\Omega \end{aligned} \tag{2.9}$$

and thus

$$D = \frac{P(\theta,\phi)_{max}}{\frac{1}{4\pi}\iint_{4\pi} P(\theta,\phi)d\Omega} = \frac{1}{\frac{1}{4\pi}\iint_{4\pi}\frac{P(\theta,\phi)}{P(\theta,\phi)_{max}}d\Omega} \tag{2.10}$$

which can also be written as

$$D = \frac{4\pi}{\iint_{4\pi} P_n(\theta,\phi)d\Omega} = \frac{4\pi}{\Omega_A} \tag{2.11}$$

in which $P_n(\theta,\phi)d\Omega = P(\theta,\phi)/P(\theta,\phi)_{max}$ is normalised power pattern and equation 2.12 is a calculated directivity from beam area Ω_A. Beam area Ω_A is a parameter for specifying three dimensional directivity of an antenna. It is the solid angle through which all of the power radiated by the antenna would stream, if $P(\theta,\phi)$ maintained its maximum value over Ω_A and was zero elsewhere. Ω_A can often be approximated by the half-power beamwidths in azimuth and elevation angels

$$\Omega_A \approx \theta_{HP}\phi_{HP}. \tag{2.12}$$

Thus, directivity is the ratio of an area of a sphere (4π sr) to the beam area Ω_A of the antenna. From the same equation, it can be seen that the smaller the beam area, the larger the directivity. The beam area $\Omega_A = 4\pi$ sr means actually *isotropic* antenna, which has the lowest possible directivity $D = 1$. All practical antennas have directivity greater than one. For example, short dipole antenna has beam area $\Omega_A = 2.67\pi$ sr and its directivity $D = 1.5 = 1.76$ dBi. dBi stands for decibels over isotropic. The antenna *gain* (G) is a realised quantity which is less than directivity because of ohmic losses. It can be written as

$$G = kD \tag{2.13}$$

in which k is efficiency factor. In other words, gain is an antenna directivity which takes also the losses into account. [13]

2.1.5 Radar Equation

Radar equation is means to determine the performance of a radar relative to its operation parameters. It can be derived with the known transmission power P_t and antenna gain G_t. G_t can be calculated by comparing the strength of the field in the direction of interest to the field caused by an isotropic antenna. Power density S_t at range R is

$$S_t = \frac{P_t G_t}{4\pi R^2}. \tag{2.14}$$

Only a small partition of the reflected signal scatters back to the radar. The area from which the target captures all the power, which it then radiates isotropically, is called *radar cross section of the target*. It is denoted with σ, and when it is taken into account, the power density returning to radar receiver is

$$S_t = \frac{P_t G_t}{4\pi R^2} \times \frac{\sigma}{4\pi R^2}. \tag{2.15}$$

Antenna captures power with the effective area A_r. Thus, the received power can be written as

$$P_r = S_r \times A_r = \frac{P_t G_t}{4\pi R^2} \times \frac{\sigma}{4\pi R^2} \times A_r. \tag{2.16}$$

If the radar is using the same antenna for transmission and reception, the gain for both is the same:

$$G_r = G_t = G. \tag{2.17}$$

By using this, the effective area A_r can be written as

$$A_r = \frac{G_r \lambda^2}{4\pi}. \tag{2.18}$$

With this the classic form of radar equation can be written as

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4}. \tag{2.19}$$

The range can be evaluated from 2.19

$$R = \sqrt[4]{\frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 P_r}}. \tag{2.20}$$

Equation 2.19 states that the received power is inversely proportional to range raised to the power of 4. This means that the strength of the point-target echo decreases rapidly as the range increases. A point-target is defined to be a target which is small

when compared with the radar resolution. [11]

2.2 Essential Theory on Sound

The word *sound* has two meanings:

1. An auditory sensation in the ear;
2. The disturbance in a medium that can cause this sensation. [14]

The first refers to an abstract and human related definition, whereas the second definition, because of its relation to physics, is of more interest here.

2.2.1 Soundwaves

Sound was previously defined as a disturbance of a medium. Air is the most common medium, but sound can also travel in any solid, liquid, or gas. Sound travels in a medium as a *sound wave*. The following wave equation presents how a wave behaves in homogenic fluid:

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 p}{\partial t^2}. \tag{2.21}$$

In the equation, x is the position, t is time, p is the sound pressure and v is the velocity of sound. The anomaly of pressure from its static value in medium causes the wave motion. This motion is moving with speed v as presented on the right side of the Equation 2.21. [15]

An example of this anomaly is the movement of a loudspeaker diaphragm, which causes periodic pattern of displacement versus time. Often this pattern is *sinusoid* or *sine wave*, which is illustrated in upper part of the Figure 2.4. In fact, sound can be thought of a combination of multiple sine waves and certain additional components. Thus, the ideal sine wave is perfect for illustration purposes. A sine wave is periodic: it repeats a certain pattern. It is a function of time, but also a function of a trigonometric angle. The formula of the sine wave is the following:

$$x(t) = A_0 \sin(\omega_0 t) + \phi_0 \tag{2.22}$$

in which A_0 is the amplitude, ω_0 is the angular velocity, t is time and ϕ_0 is the initial phase angle. The angular velocity ω can also be written as

$$\omega = 2\pi f \tag{2.23}$$

where f is frequency. It takes 2π in radians to sine wave to repeat itself; this is called a *cycle*. The time it takes for sine wave to repeat one cycle is called a *period* and is denoted by T. The distance a wave travels in one period is called *wavelength*

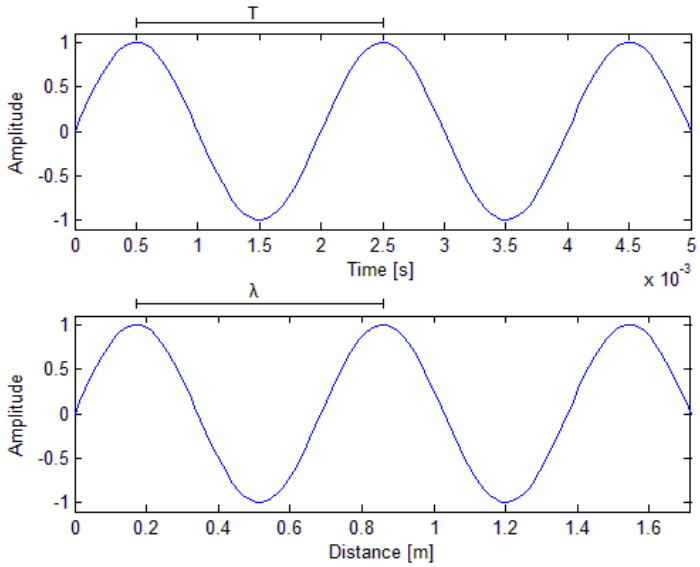

Figure 2.4: Sine wave vs time and distance.

and is is denoted by λ. This is illustrated in the lower part of Figure 2.4. The relationship between wavelength and period can be written as

$$\lambda = vT \tag{2.24}$$

which can also be modified into the following way, when frequency $f = \frac{1}{T}$ is known

$$\lambda = \frac{v}{f} \Leftrightarrow f = \frac{v}{\lambda} \Leftrightarrow v = f\lambda. \tag{2.25}$$

[16; 17]

2.2.2 Propagation of Sound Waves

The propagation of sound depends heavily on the medium. Because of the way the particles are attached to each other, sound propagates faster in solid material than in gas. In solid material the velocity of sound is

$$v = \sqrt{\frac{E}{\rho}} \tag{2.26}$$

in which E is Young's modulus of the material and ρ is density of the material, whereas in gas and fluid the velocity of sound is

$$v = \sqrt{\frac{K}{\rho}} \tag{2.27}$$

in which K is Bulk modulus. [18] Equation 2.27 can be also written as

$$v = \sqrt{\frac{\kappa P}{\rho}} \tag{2.28}$$

in which κ is adiabatic exponent and P is pressure. The value of κ depends on the temperature and thus the speed of sound is also varying with the *temperature*. [19]

Sound suffers a great deal of loss during its travel through the medium. The intensity of sound is defined as *the power which sound wave carries per unit area*. Its unit is W/m^2 in SI system. Thus, attenuation of sound is also measured with intensity. [18]

The two most significant types of sound attenuation are attenuation due to *spherical spreading* and attenuation due to *absorption*. The first type is a loss due to spherical spreading, which is a nature of sound due to spreading from a localised source. If transmitted power spreads out evenly into a sphere of radius r, the intensity at this distance can be written as

$$I = \frac{power}{area} = \frac{P_T}{4\pi r^2} \tag{2.29}$$

in which P_T is transmitted power. From the equation it can be seen that the intensity decreases with range squared.

The second type is attenuation due to *absorption*. When sound travels through the air, its intensity I decreases amount of ΔI because of the absorption losses. The decrease of intensity depends both on the distance travelled and on the initial energy, and can be written as

$$\frac{dI}{I} = -\alpha dx \tag{2.30}$$

in which α is *absorption coefficient*. If α does not vary along the sound path, then

$$I = I_0 e^{-\alpha x}. \tag{2.31}$$

The absorption coefficient is a combination of the classical absorption and molecular absorption. The classical absorption occurs when every small volume of air is compressed and stretched by the sound pressure along the direction of propagation. This causes the shape to change or shear and this change is resisted by viscous forces. The molecular absorption occurs due to the transfer of a molecule's energy out of the translational motion and into the vibration or rotation of the molecule. [20]

2.2.3 Wave Reflection

When a wave hits the boundary of its medium, all or part of it is *reflected*. The reflected sound is called an *echo*. However, it is rare for a wave to be completely reflected. More commonly, only a part of the wave is transmitted to the second material: this is called *refraction*.

The law of reflection and refraction states the following two things:

1. The incident, reflected and refracted rays and the normal to the surface lie in the same plane.
2. The angle of incident θ_a is equal to the angle of reflection θ_r for any material and for any wavelength.

The law of reflection and refraction is valid for all types of waves. Depending on the surface, there are two types of reflection. If the surface is smooth, then the reflection angle is definitive and reflection is *specular*. On a rough surface, reflection is scattered and is called *diffuse* reflection. Both cases are presented in Figure 2.5. [21]

2.3 Sonic Detection and Ranging

A sodar is a device that uses sound waves for detecting targets. It is similar to radar, but microphones and speakers are used instead of receiving and transmitting antennas. Sound waves have much lower frequency than radio waves. Also, the propagation of radio waves is different from sound waves. Because of this, sodar has some advantages compared to radar, but also certain disadvantages. A common application for sodar is wind speed estimation, which is done by examining the Doppler shift caused by turbulent fluctuations that move with the wind. The amount of the frequency shift is proportional to the velocity of the scatterer in the beam direction: if the beam is directed vertically, the vertical wind speed can be calculated from the Doppler shift. The horizontal components are determined by tilting the beam by a small angle from the vertical component, to obtain two horizontally perpendicular components. In order to get the wind profile, three Doppler shifts have to be calculated. [22]

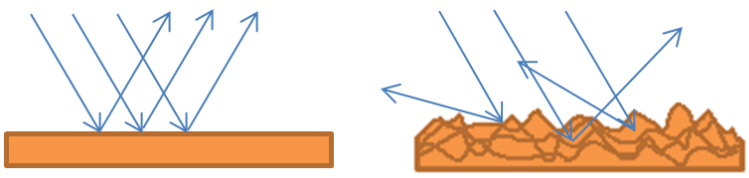

Figure 2.5: Specular and diffuse reflections.

Sodar systems are usually *monostatic*, which means that the transmitter and receiver are at the same location. It would also be possible for them to be *bistatic*, which means that the transmitter and receiver are separated spatially. A good example is a radio acoustic sounding system, which can be seen as a combination of radar and sodar. [20]

2.3.1 Sodar Equation

As was mentioned in the last section, most sodars are monostatic, because it is difficult to implement receivers and transmitters to operate at the same time. For this reason, a monostatic system usually sends pulses instead of continuous waves. As with radar, the acoustic power of monostatic sodar is spread over larger area as the range increases. Also, the sound reflected from the target spreads out: received intensity is proportional to range to the power of -4. While sodar has the same attenuation as radar, sound waves are more vulnerable to the absorption over travel distance, which leads to additional attenuation. Because sodars are usually used to measure wind profiles, the range equals the height and in sodar literature, is normally is denoted as z. As with radar, the sodar equation can be written as

$$P_r = \frac{P_t G A_e \sigma_s V}{\pi \theta^2 z^4}. \tag{2.32}$$

Scattering volume V can be written as

$$V = \pi (z\theta)^2 \frac{c\tau}{2} \tag{2.33}$$

in which $\pi(z\theta)^2$ is cross-section area of the expanding beam and $\frac{c\tau}{2}$ is the range dimension resolution. In other words, as the range increases, the volume of air mass inside the resolution element increases quadratically. Thus, when V in Equation 2.32 is substituted with 2.33, sodar equation can be written as

$$P_r = P_t G A_e \sigma_s \frac{c\tau}{2} \frac{1}{z^2}. \tag{2.34}$$

The effect of absorption also needs to be taken into account. It causes exponential fall off with distance. Sodar equation can then be written as

$$P_r = P_t G A_e \sigma_s \frac{v\tau}{2} \frac{e^{-2\alpha_s z}}{z^2} \tag{2.35}$$

in which P_r and P_t the received and transmitted power, G is the antenna gain, A_e is the effective receiving area, v is the speed of sound, τ is the pulse duration, z is the height, α_s is the absorption of the air and σ_s is the scattering cross section of the target. The main difference between the sodar equation in Equation 2.35 and the radar equation in Equation 2.19 is that the size of the target increases quadratically

as the range increases. The target in radar equation is a point target: its size will not increase as the range increases. Also, if a point target, such as a small ball, is put in the sodar target area, the received power is proportional to z^{-4}. [20]

2.4 Detection Theory

In this section, theory on detecting sound waves will be discussed. A sound wave can be perceived as combination of sinusoidal waves, which is useful in detection of sound waves. Important detection concepts include phase detection, discrete Fourier transform and matched filter.

2.4.1 Phase Detection

A common method in remote sensing systems is the determination of phase shift between the reference signal and the received wave. As the transmitted waveform is known, precise information from the signals can be gained. If the target is at range R, then the round-trip distance is $2R$. The total phase difference can be calculated as

$$\phi = -2\pi(\frac{2R}{\lambda}) \tag{2.36}$$

in which the minus sign indicates a phase delay. A typical usage of the phase detection is the calculation of the precise range shift of a moving target by detecting two consecutive pulses and calculating their phases. The range shift can be determined from the phase difference between these pulses. [10]

2.4.2 Discrete Fourier Transform

The Fourier transform maps the signal samples with the signal frequencies. There are four types of Fourier transforms: Aperiodic-Continuous (Fourier transform), Periodic-Continuous (Fourier series), Aperiodic-Discrete (discrete time Fourier transform) and Periodic-Discrete (discrete Fourier transform). In practise, the most important of these is a discrete Fourier transform (DFT), which transforms a periodic discrete-time signal into a periodic discrete-frequency signal. The result of this transform contains the frequency information of the original signal. Its formula is the following:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{\frac{-j2\pi nk}{N}}, \quad k = 0, ..., N-1. \tag{2.37}$$

Equation 2.37 means that for a discrete signal, it is possible to form a Fourier transform with discrete components. [23] This can be seen as a correlation of the signal of interest and different sine waves. In other words, this equals analysing the

frequency content component by component. [24]

The Fourier transform has numerous applications. As mentioned above, it extracts the frequency information. The result can also be expressed as the time domain data is converted into frequency domain data. Often, the frequency content itself is the main concern. However, in several cases the frequency domain data is easier to process than its time domain counterpart. In digital signal processing, digital filters are usually designed in the frequency domain.

2.4.3 Matched Filter

An important type of filter featured in detection theory is called a *matched filter*. In the previous sections it was stated that a matched filter maximises SNR. This means that the matched filter maximises the detectability of weak echo signals and attenuates unwanted signals. [4] A filter is a matched filter for a certain signal when its impulse response is a *time-inverted complex conjugate* of the signal. [25] If the received pulse shape is $h(t)$, the receiving filter is a matched filter when its impulse response is $h^*(-t)$. This is illustrated in Figure 2.6. Fourier transform of this pulse is $H(f)$: the equivalent for matched filter is then $H^*(f)$. This is an important property because

$$|H^*(f)| = |H(f)|. \tag{2.38}$$

Thus, a matched filter emphasises frequencies which are strong in the spectrum, and attenuates weaker frequencies. Since it aims to have zero-response outside the signal band, it filters noise outside the signal spectrum. Matched filter also works as a phase corrector because overall transfer function

$$H(f)H^*(f) = |H(f)|^2 \tag{2.39}$$

is real. [26]

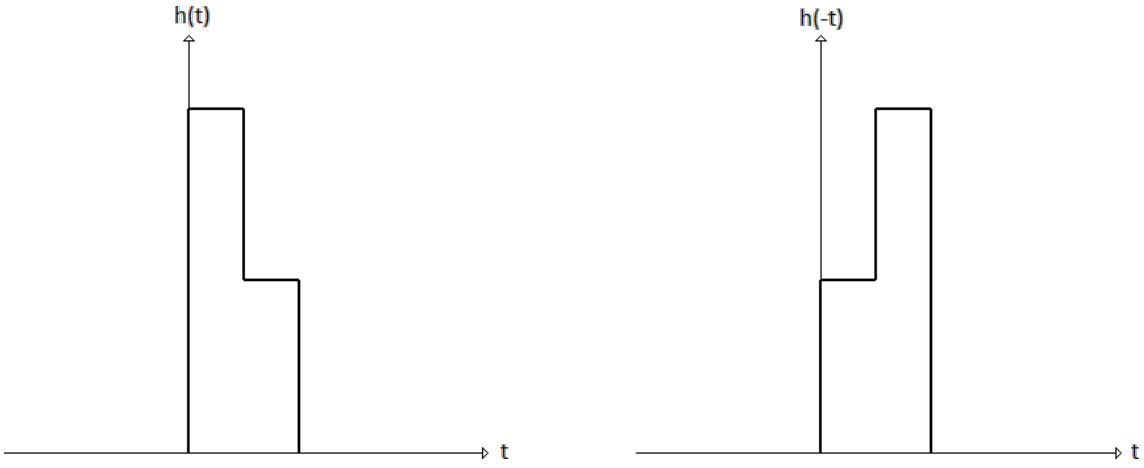

Figure 2.6: Illustration of pulse and its time reversed counterpart.

3. ARRAY THEORY

For several decades, sensors have been widely used in different kinds of applications. In addition to radar, sonar and sodar, there are several other sound and communication systems. Antennas, microphones and loudspeakers are common examples of sensors. As was stated in preceding chapter, a single sensor has limited means for directivity. Certain changes in the physical structure can be made to improve the directivity. However, with a certain physical structure, the directivity of antenna is fixed: thus, the improvement gained by these changes is rather limited. A more effective method is needed. An effective approach to improve the directivity is to use multiple sensors and to form an array. The following chapter will discuss sensor array theory. The chapter is organised in the following way; Section 3.1 explains the term *spatial filtering*. Section 3.2 focuses on the concept of *beamforming*, starting from basics and ending with a case which includes broad range of frequencies. In Section 3.3, the *window functions*, used in digital signal processing, are discussed briefly. In Section 3.4, the cases of *non-omnidirectional* sensors and *non-linear array geometries* are discussed.

3.1 Spatial Filtering

The Oxford Dictionary of English defines the word *spatial* as "relating to space" [27]. This is a valid definition in spatial signal processing as well: it means that the processing takes place in spatial domain. In other words, the subject of interest is the function of a position or direction [28]. Spatial filtering is used in many fields of science, such as geography and digital image processing [29; 30]. One of the oldest uses of spatial filtering is collecting spatial samples of wave fields which were used to form a pencil beam. Hence, the technique is called *beamforming*. It should be noted that the term also has a broader meaning: it refers to discriminating between signals, based on the physical locations of the signal sources. [3] In this thesis, beamforming is the relevant form of spatial filtering and thus spatial filtering is used to describe beamforming.

3.2 Beamforming

The beamformer is a device for forming a desired directivity or beam pattern. In the case of sensors, a beamformer can be made by placing sensors together and forming

a sensor array. First, the concept of beamforming will be discussed with uniform linear arrays (ULA). This means that the sensors are on the same line, and spacing is constant between all the elements. This is depicted in Figure 3.1. All the following cases make an assumption that source is considered to come from *far-field*. This means that wavefronts arriving at the aperture can be considered as plane waves and curvatures can be disregarded. This assumption is valid if

$$|r| > \frac{2L^2}{\lambda} \tag{3.1}$$

in which r is distance from the array, L is the effective length of the array and λ is wavelength. Waves coming from far-field and near-field are depicted in Figure 3.2. [31]

In order to use sensor array for improving directivity, all the elements of the array should receive the signal. These outputs are weighted with complex weights and then summed to combine the gain of all elements. The idea is similar as in FIR-filtering, only in spatial domain; this method is called *filter-and-sum beamforming* (FS). If the incoming signal is coming precisely from the direction to which the array is pointing at, then the signal arrives to every element simultaneously. Summing these in-phase signals cause gain to be amplified by the number of elements in the array. This is illustrated in Figure 3.3. If the signal is coming from some other direction, then it reaches every sensor at the different point of time. In this situation, the signals are out of phase and depending on the difference of the signal phases, the result of the sum will be either constructive or destructive. When this is done for an angle band, such as, from -90° to 90°, the result is *spatial spectrum*, also called a *beam pattern*. This [-90°,90°] band is called a *visible region*. Equation 3.2 describes a beam pattern for ULA with gain normalised to unity at broadside:

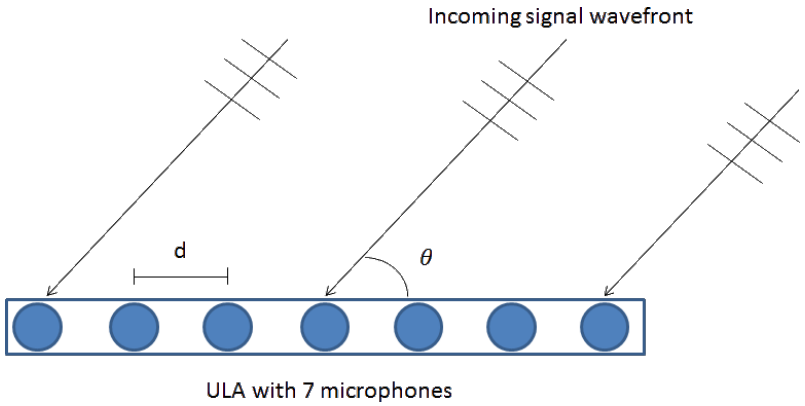

Figure 3.1: Uniform linear array with 7 sensors.

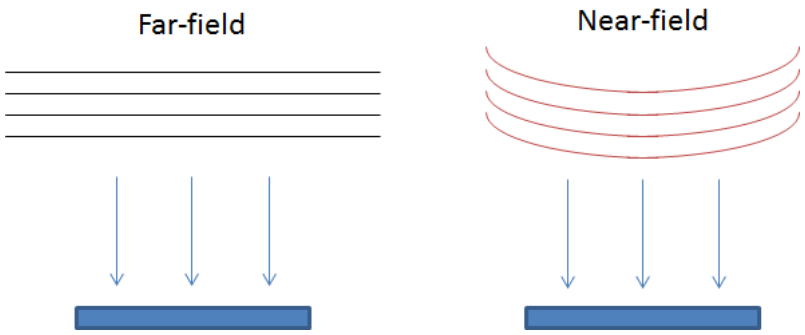

Figure 3.2: Signals coming to linear array aperture from far-field and near-field source.

$$E_a(\theta) = \frac{\sin[N\pi(d/\lambda)\sin\theta}{N\sin[\pi(d/\lambda)\sin\theta]} \tag{3.2}$$

in which, N is the number of sensors, d is the element spacing, λ is the wavelength and θ is the incident angle. $E_a(\theta)$ is also known as *array factor*. Theoretical beam pattern, with 8 kHz sine wave as a signal, is depicted in Figure 3.4. [32; 33]

The main factors which affect the beam pattern are spacing of elements, size of the array and signal frequency. The element spacing is the key factor in avoiding *spatial aliasing*. If the spacing between the elements is too wide, then *grating lobes* are encountered. They are beams which have the same magnitude as the main beam, and repeat periodically. To avoid grating lobes within visible region, the element spacing d has to be:

$$d < \frac{\lambda}{2} \tag{3.3}$$

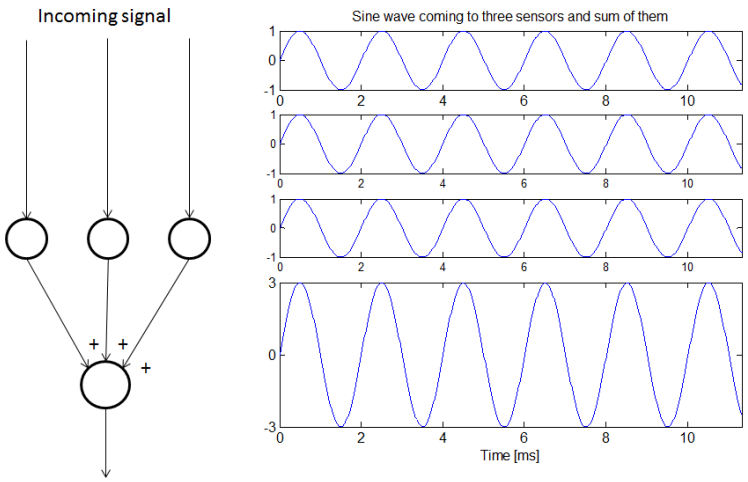

Figure 3.3: Illustration of beamformer with three input signals and an output signal.

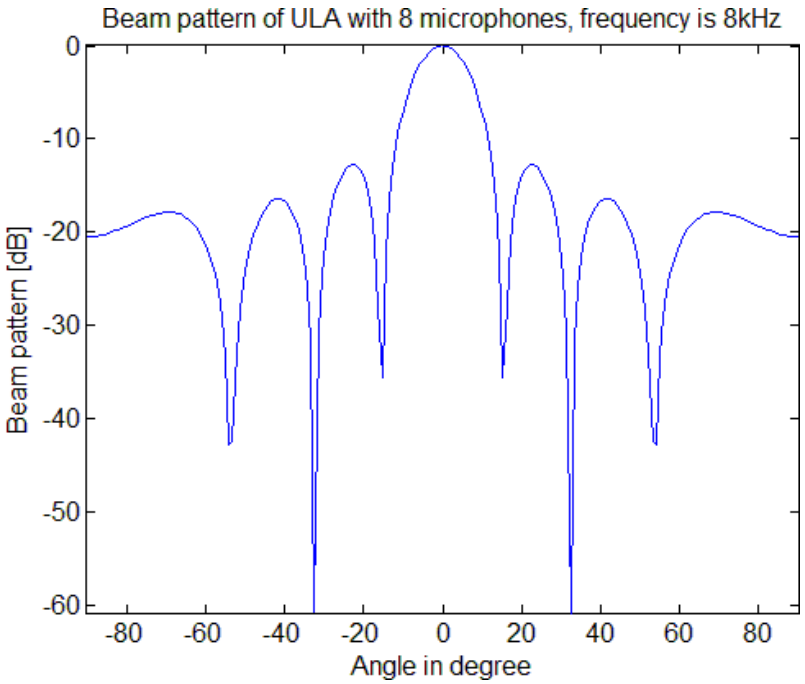

Figure 3.4: Beam pattern of ULA with 8 microphones. Signal is sine wave with frequency of 8 kHz and spacing is 2 cm.

in which d is the element spacing and λ is the wavelength. This is called a *spatial sampling theorem*. [34] A situation when this condition is not valid is depicted in Figure 3.5. The length of the array, or the size of the aperture, is more related to array's ability to discriminate between two separate incoming waves from different directions. In other words, the bigger the aperture is, the narrower the main beam. Because of the Equation 3.3, the only way to widen the array is to add more sensors. The beam pattern also depends on the frequency. It can be seen from Equation 3.3 that the maximum element spacing depends on the wavelength. Wavelength is inversely proportional to frequency. This means that the higher the frequency, the smaller element spacing we need, and vice versa. If the element spacing and the amount of elements are fixed, the main beam will be wider on the lower frequencies and narrower on the higher frequencies. In order to avoid grating lobes, the spacing has to be chosen according to the demand of highest frequency component. In Figure 3.6 beam pattern of the same ULA as in Figure 3.4 is illustrated, but this time the frequency is from 2 kHz to 8 kHz. [35]

Grating lobes, which were introduced above, are an important parameter to take into account when designing sensor arrays. Other important parameters include *3-dB beamwidth*, *height of first sidelobe*, *locations of nulls* and *locations and rate of decrease of sidelobes*. The half-power beamwidth (also called a 3-dB beamwidth) is the measure of the width of the main beam. It is defined as a point where the gain is the half of the beam maximum, and it can be calculated as

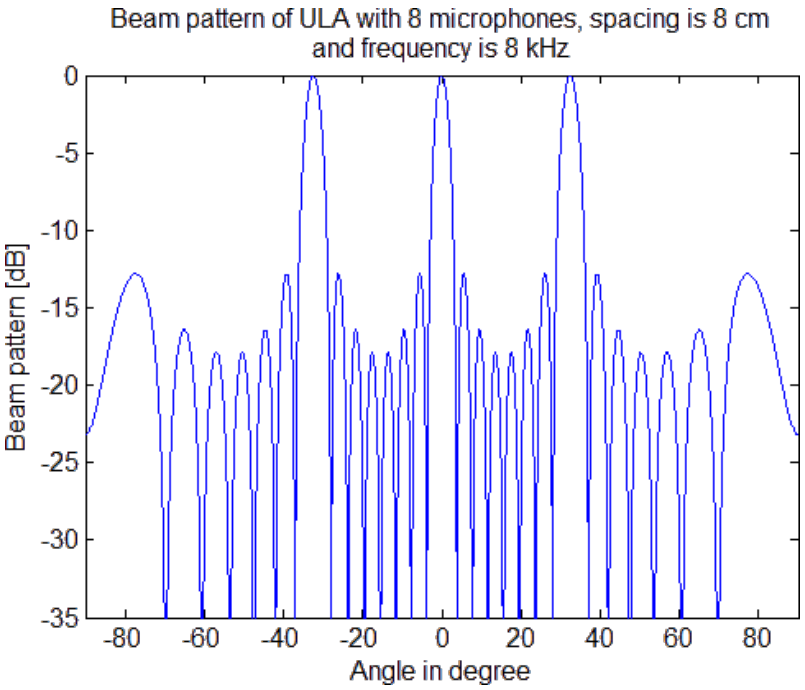

Figure 3.5: Beam pattern of ULA with 8 microphones. Signal is sine wave with frequency of 8 kHz. Spacing is 8 cm and two grating lobes can be seen at around $\pm 30°$.

$$HPBW = 2\sin^{-1}(0.446\frac{\lambda}{Nd}) \tag{3.4}$$

for arbitrary d which can be written as

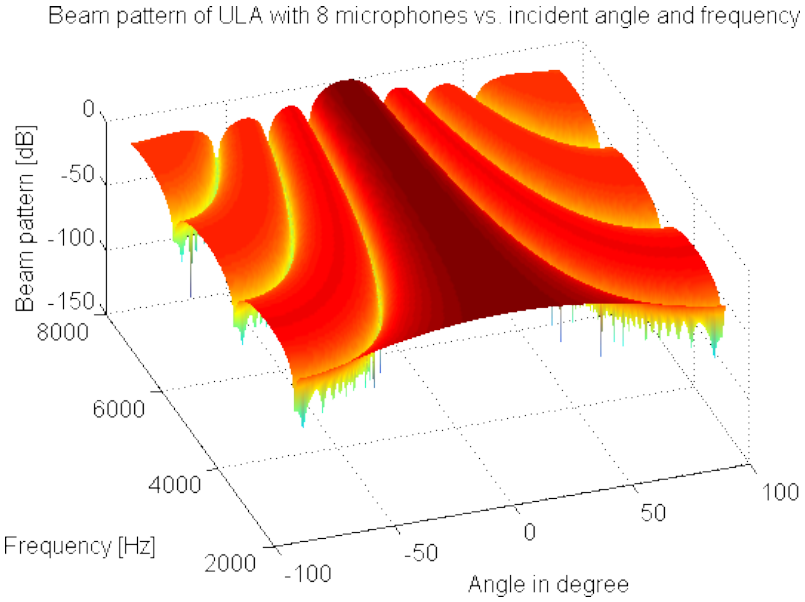

Figure 3.6: Beam pattern of ULA with 8 microphones. Signal is sine wave with frequency from 2 kHz to 8 kHz.

$$HPBW = 2\sin^{-1}(0.891\frac{1}{N}) \tag{3.5}$$

if d is half of the wavelength. Nulls are points where the signal is nearly cancelled. The nulls on the beam pattern occur when the numerator of E_a is zero and the denominator is non-zero. Thus, nulls occur when

$$\theta = \pm\sin^{-1}(m\frac{\lambda}{Nd}), \quad m = 1, 2, \ldots \tag{3.6}$$

and

$$\theta \neq \pm\sin^{-1}(m\frac{\lambda}{d}), \quad m = 1, 2, \ldots \tag{3.7}$$

Sidelobes maxima occur approximately when numerator in Equation 3.2 is a maximum. This is the case when

$$\theta = \pm\sin^{-1}(\frac{2m+1}{2Nd}\lambda), \quad m = 1, 2, \ldots \tag{3.8}$$

and accordingly, the first side lobe is located at

$$\theta_{s1} = \pm\sin^{-1}(\frac{3\lambda}{2Nd}). \tag{3.9}$$

At the maximum, the numerator is approximately one; thus, the maximum value for first side lobe can be calculated with

$$E_a(\theta_{s1}) \cong \frac{1}{N\sin(\frac{3\pi}{2N})} \tag{3.10}$$

and according to this, the sidelobe level for first sidelobe with 8 sensors is about -13 dB. The second is located approximately at θ is 42° and the sidelobe level is about -17.9 dB. These sidelobe levels are rather high, and thus a method for controlling the side lobes is necessary. Because the sensor arrays work in the same way as FIR filters, *window functions* can be used for this purpose. [36]

3.2.1 Delay-and-Sum Beamforming

There are two fundamental ways to steer a beam to a desired direction: mechanical and digital beam steering. In the old fashioned mechanical steering, the whole array is steered mechanically towards the direction of interest. Thus, the direction of the beam is decided by the actual direction to which the sensor array is physically pointing to. In digital beam steering, the array is pointing to one direction, but the complex filter weights are designed so that the beam is turned towards the direction of interest. Traditionally, this is accomplished by weighting the sensors with delay. This means that each sensor receives the signal at slightly different

point of time. This causes the beam to steer towards the desired direction.This type of beamformer is called the *delay-and-sum beamformer* (DS). A DS beamformer is illustrated in Figure 3.7. The ouput of the DS beamformer can be written as:

$$b(t) = \sum_{n=1}^{N} \omega_n x_n(t + \tau_n) \tag{3.11}$$

in which, τ_n is the time shift, ω_n is the filter weight and x_n is the incoming signal. In digital beam steering, τ_n is often called a *steering vector*. Time shift (or time difference of arrival), relative to sensor 1, can be calculated for ULA with the following equation for the electromagnetic wave:

$$\tau_n(\theta) = n\frac{d}{c}\sin\theta_0 \tag{3.12}$$

in which, d is the sensor spacing, c is the speed of light and θ_0 is the *steering angle*. As regards sound waves, c would be replaced with the speed of sound v. For narrowband signals, a phase shift means a delay between the microphones. From TDOA, a relative phase shift can be calculated with the equation

$$\phi_n = 2\pi f \tau_n(\theta) \tag{3.13}$$

in which f is the frequency of the signal. For ULA, the phase shift is linear: this is depicted in Figure 3.8. In the figure, the first microphone is treated as a phase reference for the others. Thus, when the array is pointing forward, there is no relative phase shift between the microphones. This is presented as the red line in Figure 3.8, and denoted with 0° angle. When the array is pointing to another direction, the relative phase shift between the adjacent microphones is constant. In other words, if the phase shift between the first and the second microphone is τ_1, then the phase shift between the first and the third is $\tau_2 = 2\tau_1$, between the first and the fourth

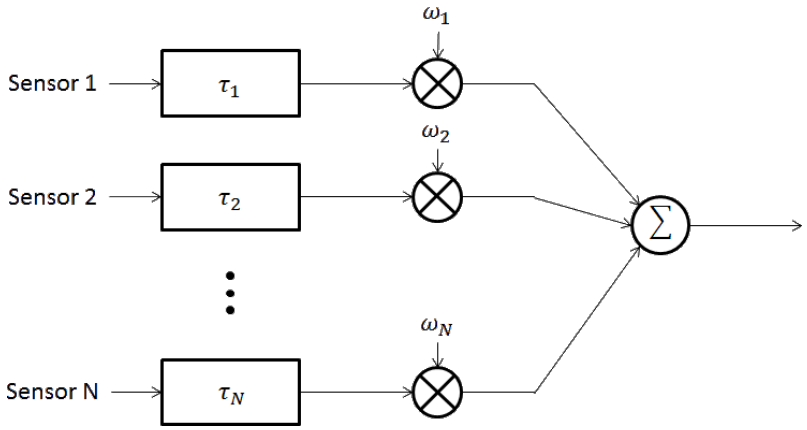

Figure 3.7: Delay-and-sum beamformer.

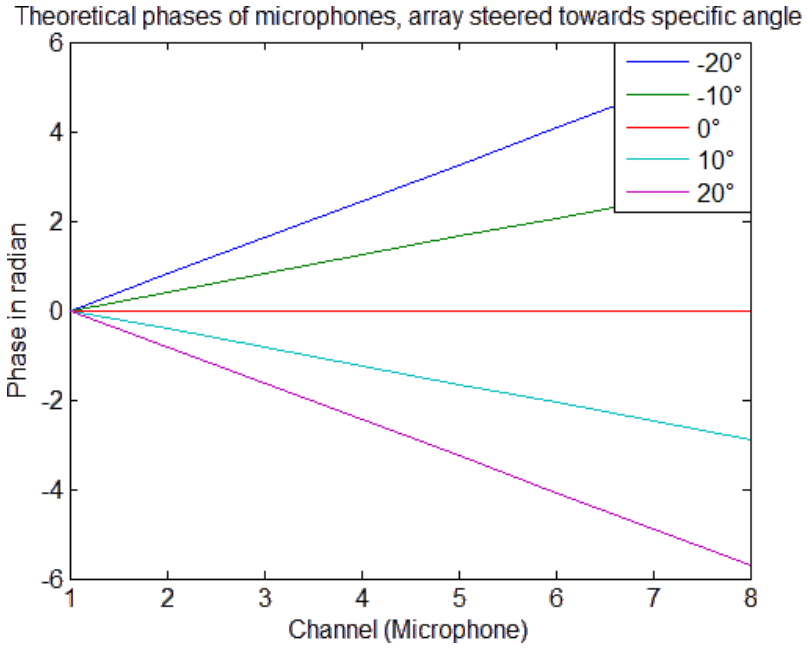

Figure 3.8: Theoretical phases of the microphones when the array is steered towards five different angles.

$\tau_3 = 3\tau_1$, and so on. [37]

The beam pattern of DS beamformer can be described with the following equation

$$E_a(\theta) = \frac{\sin[N\pi(d/\lambda)(\sin\theta - \sin\theta_0)]}{N\sin[\pi(d/\lambda).(\sin\theta - \sin\theta_0]} \tag{3.14}$$

To prevent grating lobes appearing within [-90°:90°], the following condition must be met:

$$\pi\frac{d}{\lambda}|\sin\theta - \sin\theta_0| < \pi \Leftrightarrow \frac{d}{\lambda} < \frac{1}{1 + |\sin\theta_0|} \tag{3.15}$$

which is always true when spatial sampling theorem condition in Equation 3.3 is true. [33]

3.2.2 Broadband Beamforming

The delay-and-sum beamformer is an example of broader category: the filter-and-sum beamforming. As was previously mentioned, the directivity of the sensor array changes with the frequency. This may not be a problem, for example, in the case of narrowband signals. However, incoming signals can also be *broadband*, which means that the signal frequency is in broad range. This means that the signal frequency covers multiple frequency decades. Human speech is an example of a broadband signal. There are also many applications for broadband signals depending on the system criteria. For this reason, the methods for dealing with the changing

directivity when using broadband signals are necessary. These methods form a category called a *broadband beamforming*.

One subcategory of broadband beamforming is called *constant directivity beamforming* (CDB). The aim of this method is to keep beam pattern constant over the desired frequency range. Two methods for CDB will be introduced here: the use of multiple sub-arrays with different spacing, and dividing broad frequency band to narrowband chunks and designing filter-and-sum beamformer for each chunk separately. The former is the very classic way for CDB. The idea is to use multiple arrays of sensors in a way that the spacing between the elements is different in each array. A single element can be located in multiple arrays. This is called a *nested array*. The idea of nested array is illustrated in Figure 3.9. The whole idea is based on what was discussed in above: lower frequencies have longer wavelength and, thus, need a wider array. In the classical nested array, the amount of sensors is the same for each sub-array. The spacing in each sub-array is different, depending on the frequency. In other words, an array with a different width is used for a different frequency. This will, indeed, lead to an approximately constant directivity beam pattern when enough sensors are used. However, the amount of sensors might become significant. [35]

The second CDB method is to use FS beamforming paradigm for decomposing broadband beamformer into a set of narrowband beamformers. In this case, the problem is to find correct filter weight for every narrowband beamformer so that the beam pattern is constant for the whole broad band of frequencies. Decomposition can be made either in time or frequency domain. The structure of frequency domain FS beamformer is depicted in Figure 3.10. [38]

3.3 Window Functions

The window functions are an important concept in digital filter theory. They are also a useful concept for spatial filters. The beamformers that were introduced in Section 3.2 can be thought of as using a *rectangular window* on a wide antenna. A rectangular window means that all the values of the signal are multiplied with 1 within the window and with 0 outside of it. The sensor array can be thought of

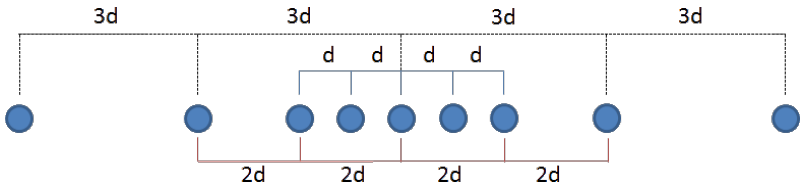

Figure 3.9: Nested array consisted of 3 arrays with 5 elements with element spacings d, 2d and 3d. Sub arrays are denoted with different coloured lines.

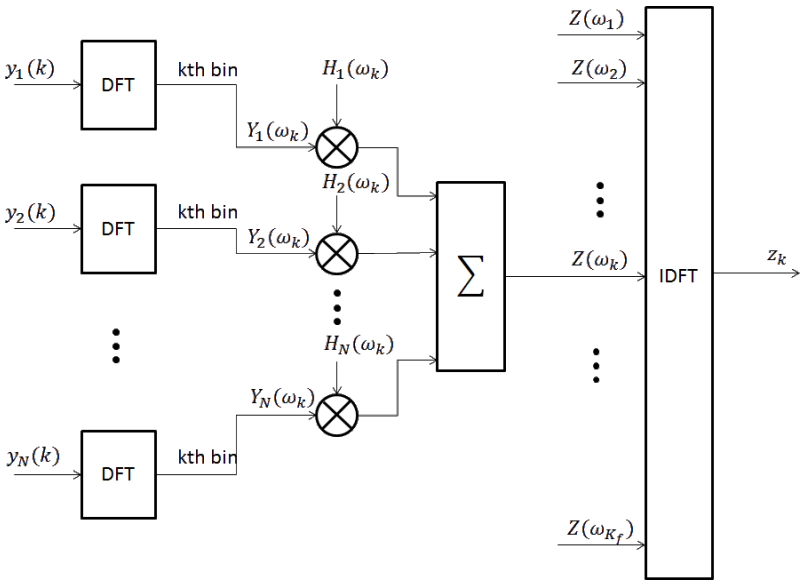

Figure 3.10: Block diagram of frequency-domain broadband beamformer. $y_{1...N}$ are sensor inputs, $Y_{1...N}(\omega_k)$ are frequency domain data of kth narrowband bin, $H_{1...N}(\omega_k)$ are filter coefficient for narrowband beamformer of kth bin, $Z(\omega_k)$ is the frequency domain output of narrowband beamformer of kth bin.

discrete version of a continuous antenna; thus, if the amplitudes are not weighted, the situation equals using a rectangular window. The Fourier transform of a rectangular window is the sinc function. A rectangular window and its Fourier transform are illustrated in Figure 3.11. The main lobe is the frequency domain characteristic of the window. The rectangular window has the narrowest possible main lobe. However, it also has high sidelobe levels: for example, the first sidelobe has -13,3 dB level. Because high sidelobes are usually unwanted, other window functions are needed. However, when the sidelobe levels decrease, the main lobe widens. Some widely used examples of window functions are *Hanning*, *Hamming* and *Dolph-Chebyshev* window functions. These functions with their Fourier transforms are depicted in Figure 3.12. As can be seen from the figure, all windows have different characteristics. Some of them are presented in Table 3.1. The use case influences the choice of the window function. In the case of sensor arrays, the narrower is the main beam, the better the capability for directivity. The height of the sidelobes should remain low, because otherwise it may be impossible to tell whether the target is inside the main beam or in one of the sidelobes. Thus, finding the best trade-off between the main beam and the sidelobes is one of the most important designing tasks with sensor arrays. [3; 39; 40]

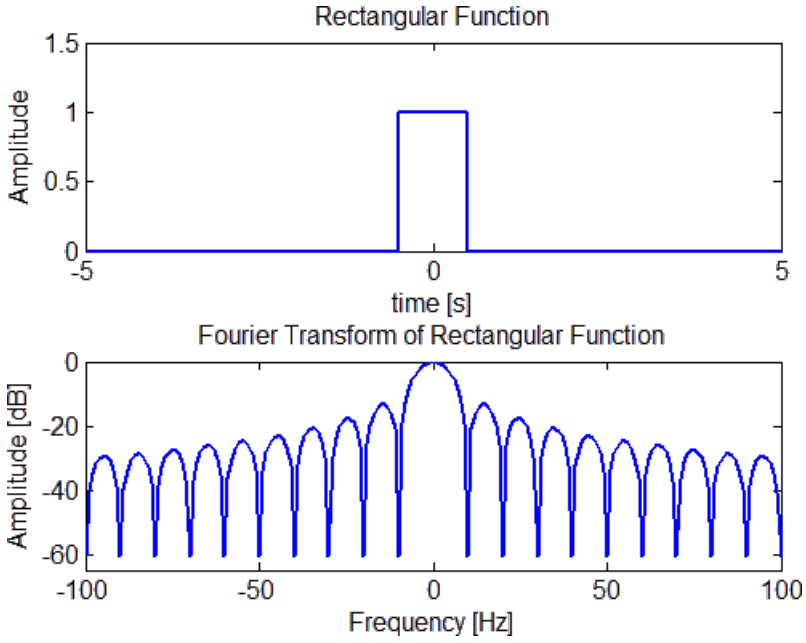

Figure 3.11: Rectangular window function and Fourier transform of it.

3.4 Non-Omnidirectional Patterns and Other Array Geometries

All the figures in preceding sections share the assumption that the used sensor is omnidirectional. This means that the sensor is sensitive to all directions equally, as illustrated in the left side of Figure 3.13. However, in practise the sensors are normally more sensitive to one direction than to others. In the case of microphones an example of the phenomenon is cardioid microphone, which is also the most common type of microphone. The vast majority of microphones belong to the *first order cardioid family*. The general polar equation for cardioid microphone is:

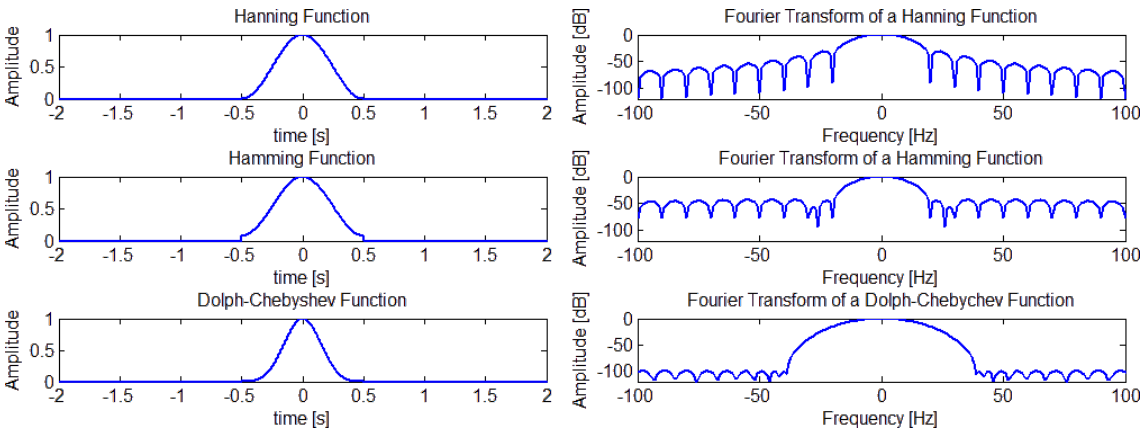

Figure 3.12: Hannig, Hamming and Dolph-Chebyshev windows.

Table 3.1: Characteristics of rectangular, Hanning, Hamming and Dolph-Chebyshev windows. [41]

Window	-3 dB Main Lobe Width (bins)	-6 dB Main Lobe Width (bins)	Highest Side Lobe Level (dB)	Side Lobe Falloff rate (dB/octave)
Rect	0.89	1.21	-13	-6
Hann	1.44	2.00	-32	-18
Hamm	1.30	1.81	-43	-6
DC, $\alpha = 3$*	1.44	2.01	-60	0
DC, $\alpha = 4$*	1.65	2.31	-80	0

*α sets Chebyshev norm of side lobes to -20α decibels

$$\rho = A + B\cos\theta \tag{3.16}$$

where $A + B = 1$. This is actually a combination of the omnidirectional pressure component (A) and the cosine gradient ($B\cos\theta$) component. If $A = 0.5$ and $B = 0.5$, then the pattern is a standard cardioid, and the equation can be written as

$$\rho = 0.5(1 + \cos\theta) = 0.5 + 0.5\cos\theta \tag{3.17}$$

which is illustrated in right side of Figure 3.13. Naturally, this also affects the beam pattern of the microphone array. The beam pattern of ULA with 8 microphones, with omnidirectional microphones and cardioid microphones is presented in Figure 3.14. As can be seen from this figure, omnidirectional microphones have identical response towards back and front. When cardioid microphones are used, the response towards back is greatly reduced.

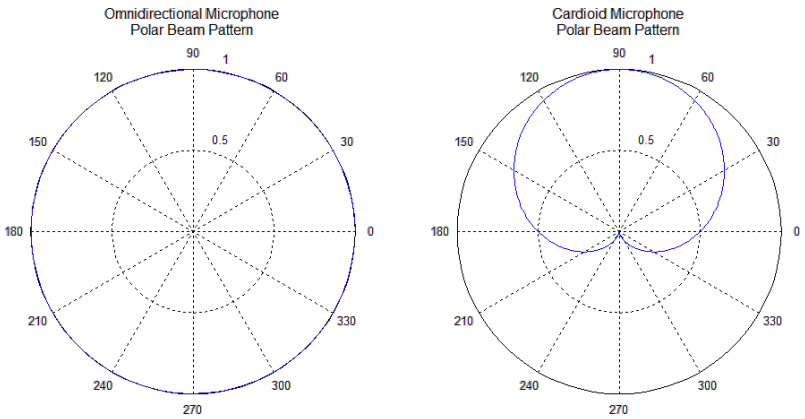

Figure 3.13: Omnidirectional and cardioid microphone beam pattern in polar coordinate system in linear scale.

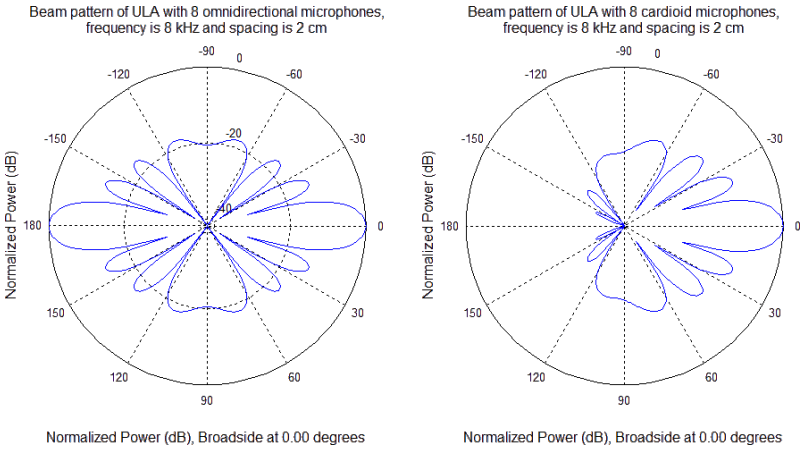

Figure 3.14: Omnidirectional and cardioid microphone array beam pattern in polar coordinate system. Broadside is at 0°.

In practise, any transmitter or receiver works, not only in two, but also in three dimensions. A linear sensor array improves the performance only in one dimension. There are also different geometrical patterns to which the sensors can be put. One of the most typical cases is a *planar array*. In this case, the array geometry is two dimensional and, thus, forms a plane. The used planar geometry depends on the requirements. A few common geometries are rectangular, circular and hexagonal. These geometries are depicted in Figure 3.15. Despite the fact that all these examples are regular geometries, the placements can be chosen arbitrarily. Expanding planar arrays to three dimensional arrays is also possible. [36]

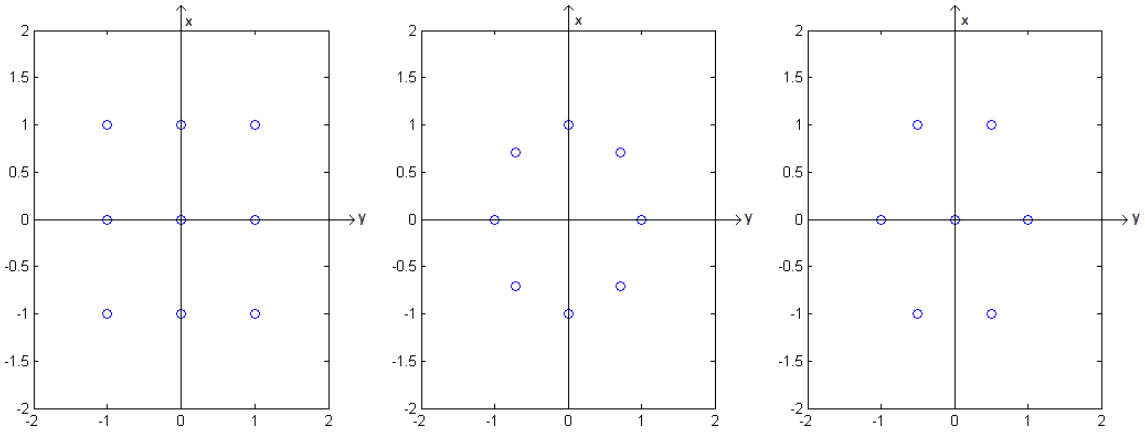

Figure 3.15: Rectangular, circular and hexagonal microphone array geometries.

4. EXPERIMENTAL SYSTEM FOR SODAR RESEARCH

In this chapter, the experimental research system, designed by Tomi Nihtilä will be discussed. Because the research system has not previously been documented in any other work, it will be documented here. The measurement methods are presented as well. The chapter consists of four sections. In Section 4.1, the overall system is described. The section documents the most important specifications of the used system components. In Section 4.2, sodar block diagram is presented and its operation is discussed. Section 4.3 describes the MATLAB environment used in the project. In Section 4.4 a microphone calibration method, which is the main contribution of the author for this research system, is presented.

4.1 System Description

The built research system consisted of an audio interface, 8 microphones, a speaker and a stepper motor. It is presented in a block diagram in Figure 4.1. The system is operated with PC and MATLAB. The mechanical array steering was done by Nanotec stepper motor, which was controlled by Arduino Uno and Arduino Motor

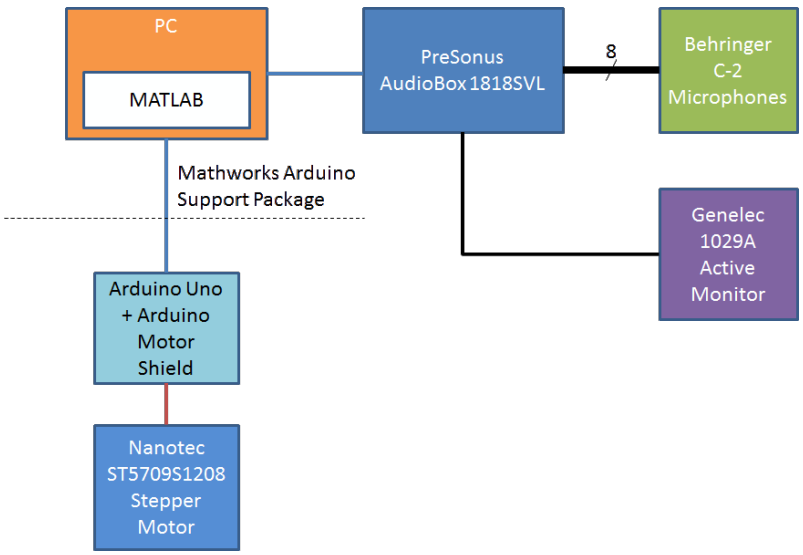

Figure 4.1: Block diagram of built sodar system.

Shield boards. Arduino boards were connected to the PC via USB port. PreSonus audio interface was connected to the PC via the USB port. Eight Behringer C-2 microphones were connected to input jacks of the audio interface. A Genelec 1029A active speaker was connected to the output jack of the audio interface.

4.1.1 Input and Output Device

PreSonus AudioBox 1818VSL was chosen as an audio interface for the project. The 1818VSL is illustrated in Figure 4.2. It is an input and output device which has 8 balanced XLR jacks for incoming signals. Each of these jacks has microphone preamplifier, and is controlled by its own input gain/trim control knob. The 48-volt phantom power for microphones was also provided. There were 8 TRS balanced line output jacks for external devices, and left and right channel jacks for the main outputs. These main outputs are controlled by a main output knob. The 1818VSL has 24 bit analog to digital converter with sampling rate up to 96 kHz. The 1818VSL can be connected with the USB 2.0 cable. Some specifications which are relevant for the thesis are presented in Table 4.1. [43]

Table 4.1: Specifications of AudioBox 1818VSL. [43]

Microphone Preamp	
Connector type	Combo, XLR, female, balanced
Mic Preamp EIN	-133 dB, 20 kHz BW, max gain, R_s = 40, A-wtd
Frequency reposense	20 Hz–22 kHz, ±0.25 dB, unity gain
Signal-to-noise ratio	97 dB, +4 dBu, 1kHz unity gain, 22 kHz BW, A-wtd
Phantom Power	+48 VDC, 64 mA total unit
Analogue Outputs	
Connector type	$^1/_4$"TRS, female, impedance balanced
Frequency reposense	20 Hz–22 kHz, ±0.25 dB
Signal-to-noise ratio	97 dB, +4 dBu, 1kHz unity gain, 22 kHz BW, A-wtd
Output Level	+10 dBu, $<$ 0.5% THD
Digital	
ADC Dynamic Range	114 dB, 48 kHz, A-wtd
DAC Dynamic Range	114 dB, 48 kHz, A-wtd
Host Interface	USB 2.0 high-speed Bit Depth
24-bit Sample Rates	44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz
Maximum Latency	6 ms, analog input to analog output

Figure 4.2: PreSonus AudioBox 1818VSL.

4.1.2 Loudspeaker

All the measurements were conducted with a Genelec 1029A Active Monitor, depicted in Figure 4.3. The 1029A is a two way active-monitoring speaker. It is designed to have high output, low coloration, and broad bandwidth. Audio input is via a 10 kOhm balanced XLR $^{1}/_{4}$" jack. Specifications of the 1029A are presented in Table

Figure 4.3: Genelec 1029A Active Monitor.

4.2. [44]

Table 4.2: Specifications of Genelec 1029 Active Monitor. [44]

Frequency response	68 Hz–20 kHz (-3 dB)
Inputs	Input 1: XLR female, balanced 10 kOhm
	Input 2: $^1/_4$" Jack socket, balanced 10 kOhm
Crossover frequency, Bass/Treble	3.3 kHz

4.1.3 Microphones

The chosen microphone model was a Behringer C-2 Studio Condenser Microphone. C-2 microphones have balanced XLR-connectors. The transducer type is a single diaphragm, true condenser type. The operating principle is the pressure gradient principle, and the polar pattern is cardioid. The frequency response is from 20 Hz to 20 kHz and the supply voltage is +48 V. This information is presented in Table 4.3 [45].

Table 4.3: Specifications of C-2 microphone.

Transducer type	True condenser, 16 mm single diaphragm
Operating principle	Pressure gradient
Polar pattern	Cardioid
Connection	Gold-plated balanced XLR connector
Frequency reposense	20 Hz–20 kHz
Signal-to-noise ratio	75 dB
Supply voltage	+48 V

The Behringer C-2 microphone is presented in Figure 4.4.

4.1.4 Stepper Motor and Stepper Motor Controller

In the initial test, the microphone array was steered by hand in order to obtain an idea of its performance. Precise measurement, however, was not possible. For this reason, a stepper motor was introduced to the system. A Nanotec ST5709S1208 was the chosen model. It has a 0.9° step angle with ±5% non-cumulative stepping accuracy. The axial force of the motor is 10 N and radial force 130 N for 5 mm distance away from the motor, 90 N for 10 mm distance, 70 N for 15 mm distance and 52 N for 20 mm distance. [46]. The stepper motor was controlled by Arduino Uno and Arduino Motor Shield. Arduino Uno is a microcontroller board based on ATmega328 from Atmel. It has 14 digital I/O pins, 6 analogue inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header and a reset

Figure 4.4: Behringer C-2 microphone.

button. [47] Arduino Motor Shield is based on L298 from STMicroelectronics. It is a dual full-bridge driver for inductive loads such as relays, solenoids, DC and stepper motors. [48] The Nanotec stepper motor, Arduino Uno and Arduino Motor Shield in use are depicted in Figure 4.5.

4.2 Sodar System Operation

The operation of the sodar is illustrated in block diagram in Figure 4.6. There are similarities with the block diagram of radar, but in the case of sodar, the whole chain is simpler, as there is no need to switch between higher and lower frequencies.

First the waveform was generated on the PC and in the Simulink environment. Then the generated signal went to the audio interface which converted the generated digital data into an analogue waveform and sent it to the speaker. The microphone array first received the transmitted signal, and as above in the transmission phase, conveyed it to the same audio interface. The audio interface converted the signal to digital and transmitted the signal back to the PC environment. In the measurements, the transmitted waveform was a sine wave. In order to to eliminate the

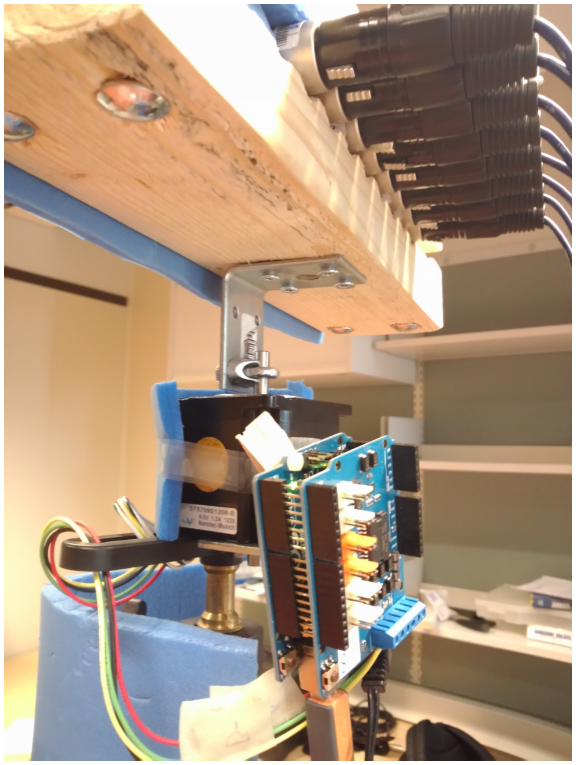

Figure 4.5: Nanotec ST5709S1208, Arduino Uno and Arduino Motor Shield in microphone array measurement system.

possible echoes, only a small pulse of the transmitted wave was used. This is depicted in Figure 4.7. In the figure, a few issues outside the actual pulse should be noted. Before the transmission begins, there is some noise from sample at around sample number 3000 until the moment of an actual transmission at around sample number 15000. This is due to the Simulink buffer with AudioBox ASIO driver: the microphones are not recording at this point. Also, the shape of recorded signal is

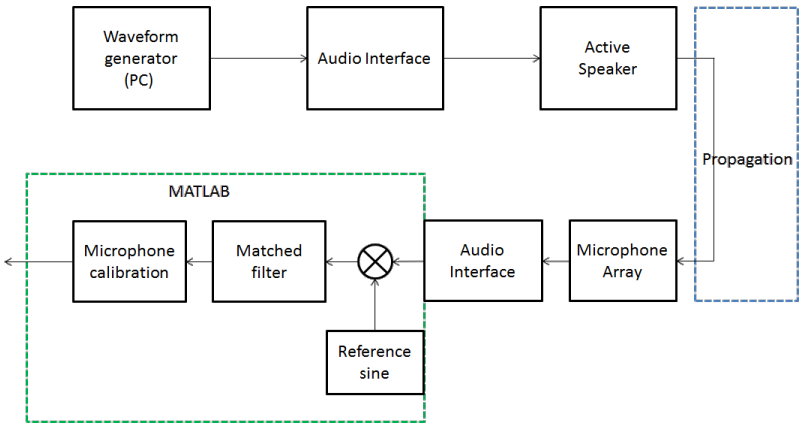

Figure 4.6: Sodar block diagram.

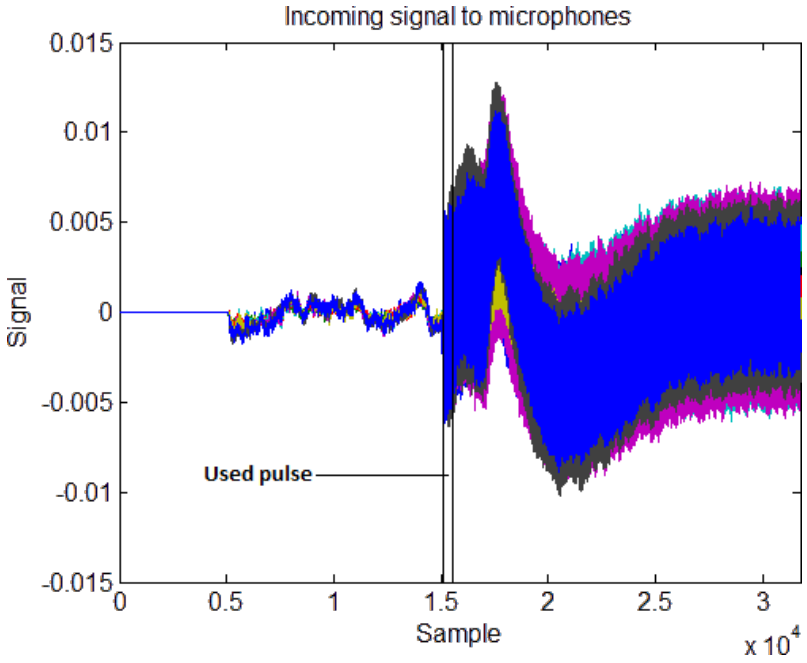

Figure 4.7: Recorded signal of the microphones. Speaker is located right in front of microphones. Used part of the signal is marked in the figure.

unusual: the shape first reaches a high value and then a low value. This tail-like shape was an unexpected microphone characteristic: the same characteristic was not visible with the reference microphones. However, since phenomenon has rather low frequency, it does not have a significant effect on the sodar performance.

The initial part of the pulse, recorded by three microphones, is presented in Figure 4.8. As can be seen from the figure, the starting point of the signal is different in each microphone. This means that, due to microphone unidealities, the signals are out of phase. It should be also noted that the shape is not a perfect sine waveform. This is due to the rather low sampling frequency relative to the signal frequency. As the original transmitted waveform is known, the received signal can be mixed with reference sine which is the same waveform as the transmitted one, except the reference sine is complex while the transmitted signal was real. The real part of the reference sine wave is depicted in Figure 4.9. The result is correlation at the frequency of the sine wave: this mixing acts in similar manner as a matched filter, but at that particular frequency. The mixing with a known waveform is known as *down-mixing* and it is used to improve SNR. The frequency response of the reference sine wave is illustrated in Figure 4.10. As can be seen, there is only a single peak in the the response. This highly ideal characteristic occurs because the sine wave has the frequency of 8 kHz and frequency resolution of fast Fourier transform (FFT) in the figure is 100 Hz. This means that the whole 8 kHz sine wave can be represented with a single frequency bin and the effect of spectral leakage is minor. When the

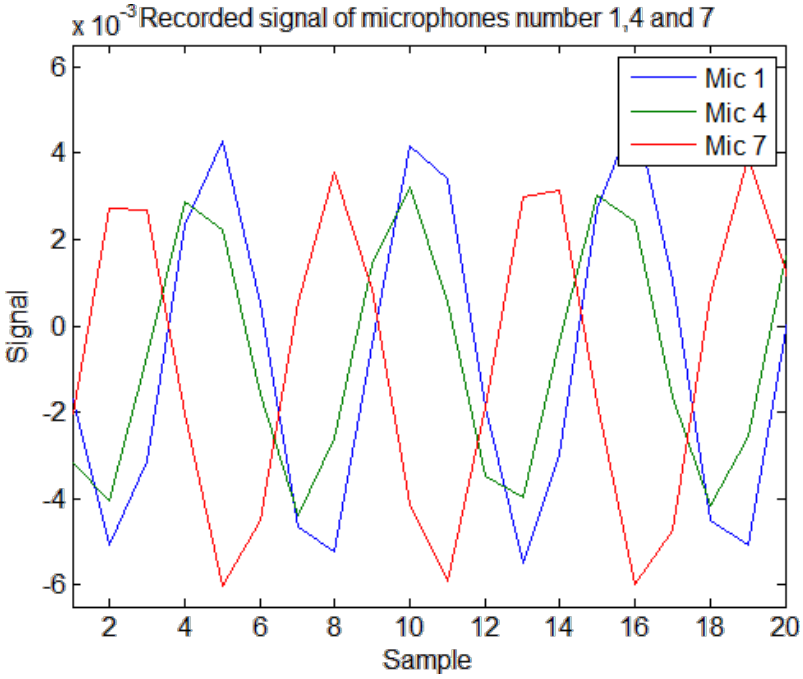

Figure 4.8: Recorded pulse from three microphones.

correlation is done in this fashion, it enables comparison of phase difference. The relative phases of pulses in Figure 4.8 relative to the reference sine are depicted in Figure 4.11. Because the speaker is located in front of the microphones, the relative phase of each microphone to the reference sine should be nearly the same. As can

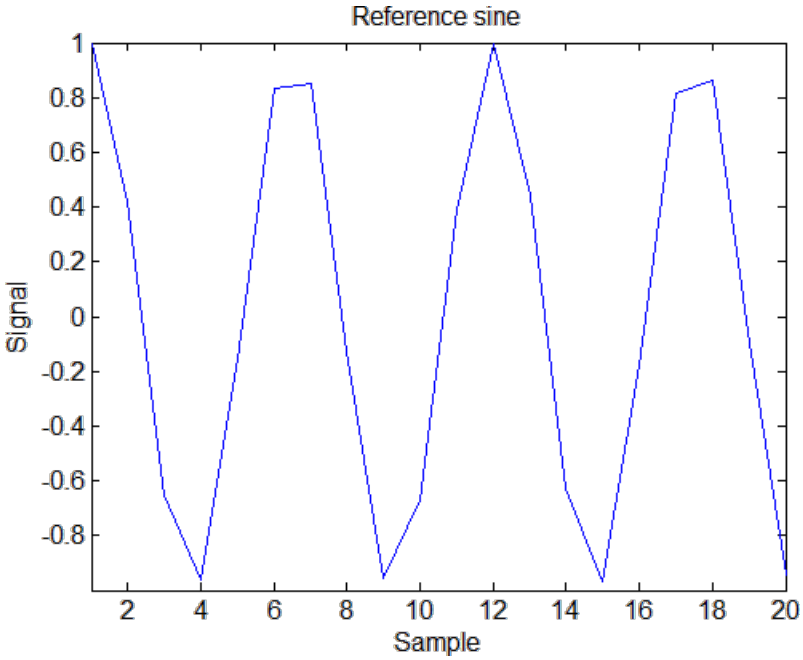

Figure 4.9: Real part of complex reference sine wave.

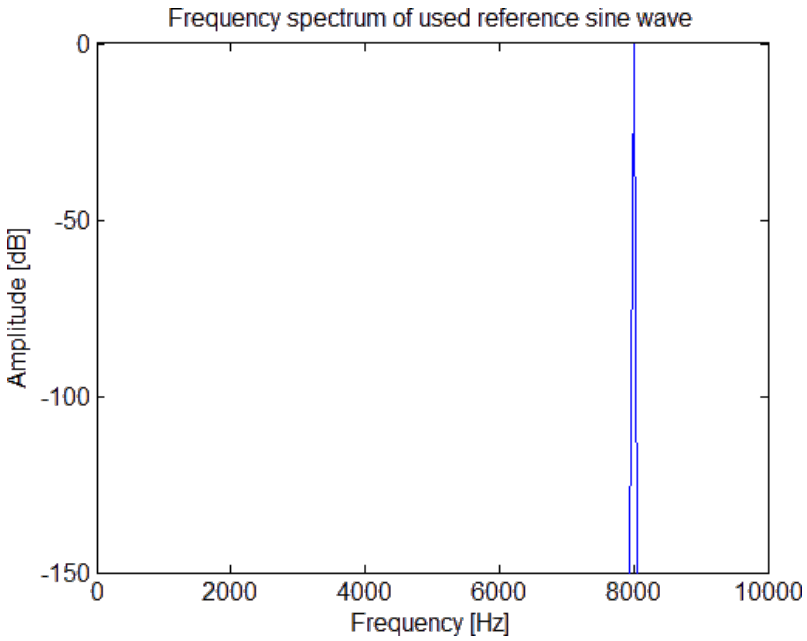

Figure 4.10: Frequency response of the reference sine wave.

be seen, the phases are quite far from each other, and thus a phase calibration is needed. In Figure 4.8, the amplitude of each microphone is slightly different and should be corrected. These calibrations are discussed in Section 4.4.

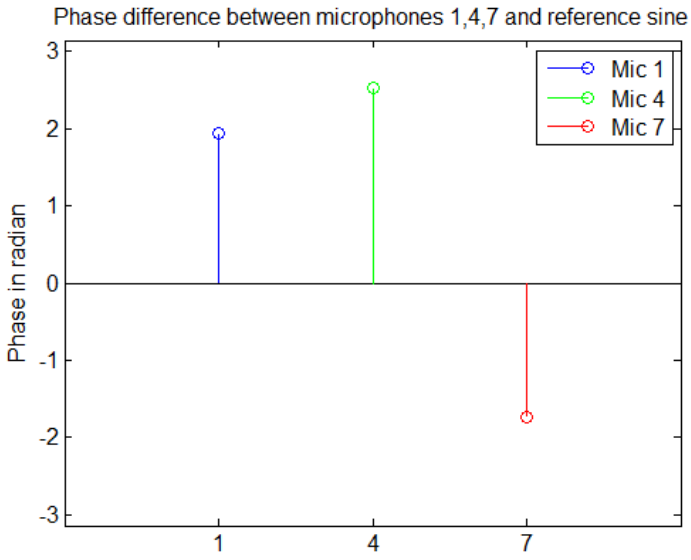

Figure 4.11: Relative phase of microphones 1,4 and 7 relative to the reference sine.

4.3 MATLAB Environment

All measurements were conducted in the MATLAB environment. The wave generation Simulink model, stepper motor control and MATLAB workspace operations will be briefly discussed in this section.

4.3.1 Sine Generation Simulink Model

A simple simulink model was used for the sine wave generation and recording. It handles the phases from waveform generation to audio interface in sodar block diagram presented in Figure 4.6. This simulink model is presented in Figure 4.12. The model has two simple tasks: a sine wave is generated and taken to speaker, and the signal is recorded with a microphone and then sent to MATLAB workspace. The sine wave is generated with the Sine Wave block of Simulink's DSP System Toolbox. The generated sine wave was chosen to be real, because all signals in nature are real signals. The Sine Wave block is connected to the To Speaker block, which takes the signal to speaker by using PreSonus AudioBox. The AudioBox is accessed straight from Simulink by using AudioBox ASIO Driver. This part is represented in the upper branch of the Figure 4.12. The lower branch of the same figure represents the recording of the signal and taking it to the MATLAB workspace. The From Mic block takes the signal from the microphone. It also uses AudioBox ASIO Driver. The signal is then taken to the MATLAB workspace.

4.3.2 Stepper Motor Control

The stepper motor used for the measurements was controlled by Arduino Uno and Arduino Motor Shield. It can be controlled directly with MATLAB via Mathworks Arduino Support Package. Every connection of both Arduino boards and stepper motor, as well as the MATLAB object to operate the motor were implemented by

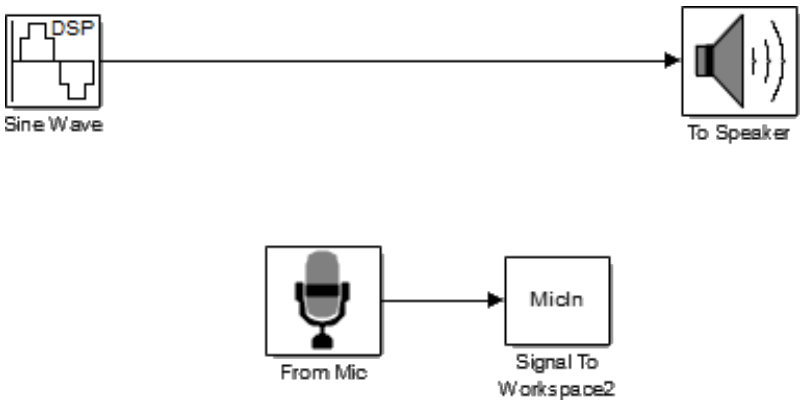

Figure 4.12: Simulink model for waveform generation and recoding the signal.

Tomi Nihtilä. The use of the motor was simple: either the number of steps, or the desired angle toward which the shaft is needed to be turned, could be entered. The motor shaft would then be turned towards the given point. Unfortunately, the stepper motor does not have feedback, which means that the initial angle has to be entered manually. The reference angle can be changed later if necessary. The shaft rotating speed may also be changed.

4.3.3 MATLAB Workspace Operations

Initiations of simulink and stepper motor was done in the MATLAB workspace. In addition, the rest of operations of sodar operation chain were also done in MATLAB. All the signal processing operations for the array, as well as the array calibrations and result plotting, were done in MATLAB. Both the simulations and measurements were done with the same operations, with the exception that the measurements used data from the microphones, and simulations were done with theoretical sine waves. At this point, the most important part was the mixing with reference sine, which was explained above in the Section 4.2. There were two kinds of measurements: calibration value measurements and actual pattern measurements. If the calibration measurement is in action, then the values are merely saved. If the beam pattern measurement was in action, then the detected and filtered signal was calibrated with calibration value in the way described in the Section 4.4.

4.4 Microphone Calibration

Microphones do not have identical characteristics, and this is even more true in the case of inexpensive microphones. There are multiple possible reasons for the phenomenon. The manufacturing process might not be precise, and some displacements of the components might exist. Also, the components are, naturally, not identical. This applies especially to cheaper microphones because the component materials are of the lower quality. This affects both the phase and the amplitude of the microphone; thus, these factors have to be compensated.

The compensation can be done by performing a calibration measurement before the actual measurement. In the calibration measurement, the same sine wave is generated as in the actual measurement. If digital steering is used, this has to be done for every incident angle, and for every frequency if the transmitted signal is broadband. This way, the phase calibration and the amplitude compensation values for every incident angle and for every frequency are attained. The phase calibration value is calculated by using one microphone as a reference. In the conducted measurements, the first microphone was taken as the reference, leading to a situation similar as in Figure 3.8. In other words, the phase calibration value is the phase

difference relative to first microphone of the array. The amplitude compensation value is attained in the similar manner. It is an amplitude relation between the chosen reference microphone and the rest of the microphones. In all the figures in the next chapter, the microphone that has the highest amplitude value was chosen as the reference. This way, a three dimensional calibration matrix was formed. In this three dimensional matrix, the incident angle is only in azimuth direction. If elevation needs to be also taken into account, then measurement needs also to be done for every elevation angle. This would result in a fourth dimension in the calibration matrix. The three dimensional calibration matrix is illustrated in Figure 4.13.

In the matrix, one dimension is the *calibration vector* which can be denoted as $A_i e^{j\theta_i}$. It includes the phase and the amplitude correction coefficients for each microphone, and the phase shift or delay needed for digital steering towards different angles. This is an efficient way to obtain both microphone errors and the steering vector with only one calibration measurement. The second dimension of the matrix is the channel, or, in this case, a microphone. The last dimension is the signal frequency. As an overall result, both the microphone calibration and steering values for multiple different frequencies can be obtained s with a single measurement.

The actual use of the attained calibration matrix is simple. The measurement value of the channel i is multiplied with the corresponding calibration value. The correspondence includes the correct channel, direction and frequency. For example, a calibration measurement with 8 microphones was done in a frequency range from 4 kHz to 8 kHz, and for the incident angles from -20° to 20°. If the uncalibrated

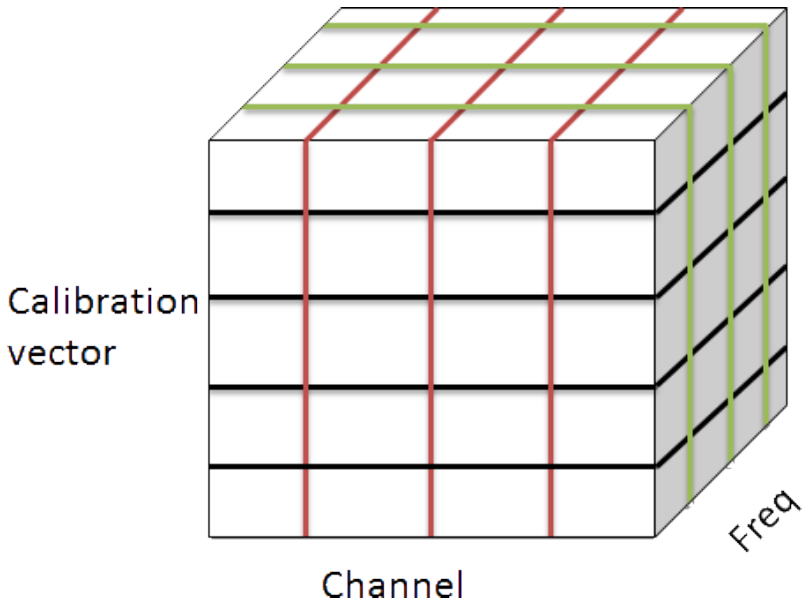

Figure 4.13: Illustration of microphone calibration matrix.

measurement is 6 kHz signal coming from 0° angle, then all 8 microphones are multiplied with 8 calibration values with 6 kHz which corresponds 0° incident angle.

5. MICROPHONE ARRAY SIMULATIONS AND MEASUREMENTS

To obtain a clear apprehension of the research cases, a theoretical simulation of each different measurement case was conducted before the actual measurements. It is important to first test the intended method with ideal cases. This way, the residual case can be extracted and used to conclude which are the effects of microphone unidealities. Thus, all the cases measured with C-2 microphones were first simulated in MATLAB. The simulations were verified by using the equations from Chapter 3. The actual measurements were done after the simulations. In all of the measurements, a speaker was put in front of the array and directed towards it, after which the array was steered, either mechanically or digitally. Figure 5.7 depicts two cases in which array is pointing towards microphone, but beam is digitally steered towards other direction. All measurements were conducted in an office environment. The location of the speaker and the microphones were chosen so that they would be as close to center of the room as possible. The idea was to get the surfaces which cause the echoes as far from the measuring system as possible. Also, only a small part of the transmitted signal was used. This way, the used partition could be chosen so that no echoes reach the speaker by the end of the measurement time-window. It should be noted that this is not an actual sodar operation, as sodar should be detecting echoes. However, the main objective was to measure the operation of the microphones. Placing the speaker in front of the array and, thus, avoiding echoes, help to clarify the measurements.

Three different research cases are covered in the following sections:

1. Simulations and measurements of mechanical steering
2. Simulations and measurements of digital steering
3. Simulations and measurements of broadband beamforming

The first case presents a mechanical steering of the microphone array. It includes the initial measurement without the stepper motor, and multiple measurements with the stepper motor. The stepper motor measurements are: a measurement without calibration, a measurement with the phase calibration, and a measurement with the both amplitude and phase calibrations. The second case presents a measurement

to verify the use of the delay for steering the beam. It also shows an actual scanning measurement with digital steering. The final case shows the simulations and measurements of broadband beamforming.

5.1 Mechanical Steering of 8 Microphone Uniform Linear Array

Because the audio interface in the project had eight input channels, the same number of microphones was used in the array. The diameter of C-2 microphone is about 2 cm: this became the smallest spacing that could be used. According to Equation 3.3, the maximum frequency that can be used is

$$d < \frac{\lambda}{2} \Leftrightarrow d < \frac{v}{2f} \Leftrightarrow f < \frac{v}{2d}. \tag{5.1}$$

When $v \approx 343$ m/s and $d = 2$ cm, the maximum frequency is approximately 8.5 kHz. For this reason, the highest used frequency in the measurements was 8 kHz. The theoretical beam pattern of this array has already been illustrated in Figure 3.4. As the array structure is fixed, the beam pattern becomes wider with the lower frequency. The simulated beam patterns with frequencies 6 kHz and 4 kHz are presented in Figure 5.1.

The first version of the measurement was done with a relatively primitive setup. The array was built of cardboard and tape. It was put on a camera stand and steered by hand. The beam pattern of 6.5 kHz sine waves with manual steering is illustrated in Figure 5.2. There are several points to be observed in the figure. First of all, there are no means to ascertain that the steering speed remains constant. Thus,

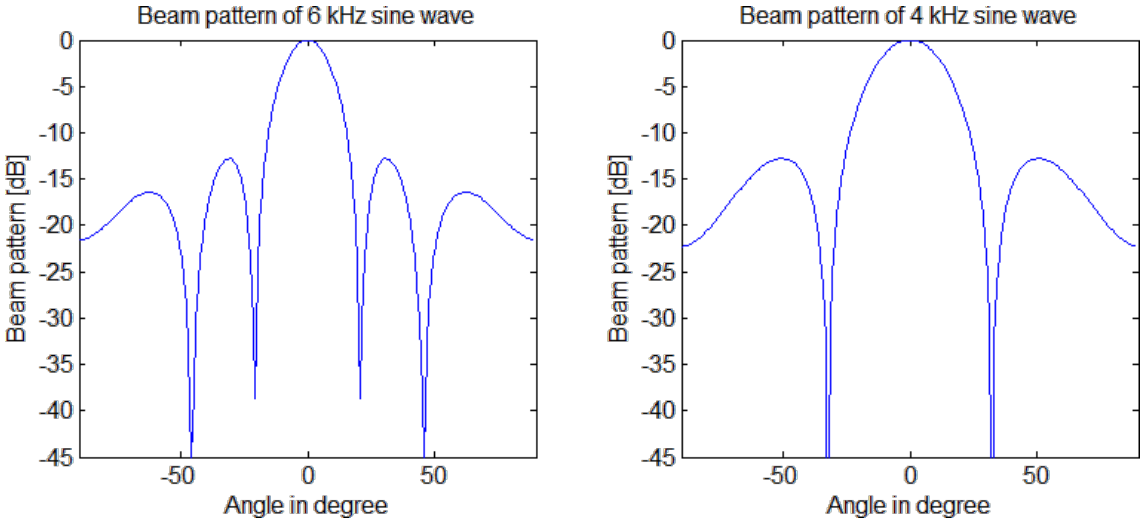

Figure 5.1: Theoretical beam patterns of 4 kHz and 6 kHz sine waves.

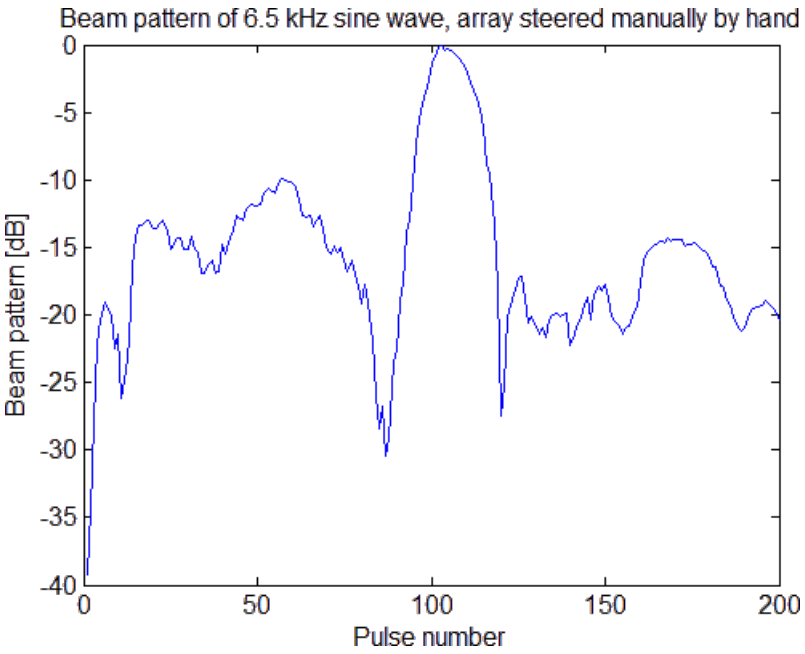

Figure 5.2: Measurement of microphone array steered manually by hand. Transmitted signal is 6.5 kHz sine wave.

the x-axis label is the transmitted pulse number, not the angle. The main lobe is distinct, but sidelobes are not that clear. Yet, the shape of the figure resembles the theoretical model. In other words, the initial measurement suggests that C-2 microphone could be capable for the system.

However, more precise measurements were needed. Tomi Nihtilä's stepper motor and stepper motor control were deployed. The stepper motor allows to make the measurements with a controlled angle setup, when the step size between transmitted signals remains the same for the whole measurement. Figure 5.3 illustrates the measured response of 8 kHz sine wave with stepper motor. The beam pattern begins to demonstrate some of expected characteristics but is still lacking in the others. Side lobes are more visible, but their levels remain far from the ideal levels, as can seen from the first side lobe on the minus degree side and from the high lobes at -80° and 80°. The phenomena are caused by the differences in microphones and can be corrected with calibration. In Figure 3.8 the theoretical phases for ULA were presented. Figure 5.4 depicts the measured phases. On the left side of the figure, the phases appear as they were measured, and on the right they are calibrated with zero-angle phases (array pointing towards speaker). As can be seen from the figure, there are tangible differences in the phases of different microphones. The calibrated figure demonstrates that with calibration these phases can be corrected towards their theoretical values. Somewhat surprisingly, the zero-angle calibration gives good results for other angles as well. There is, however, an obvious explanation. As

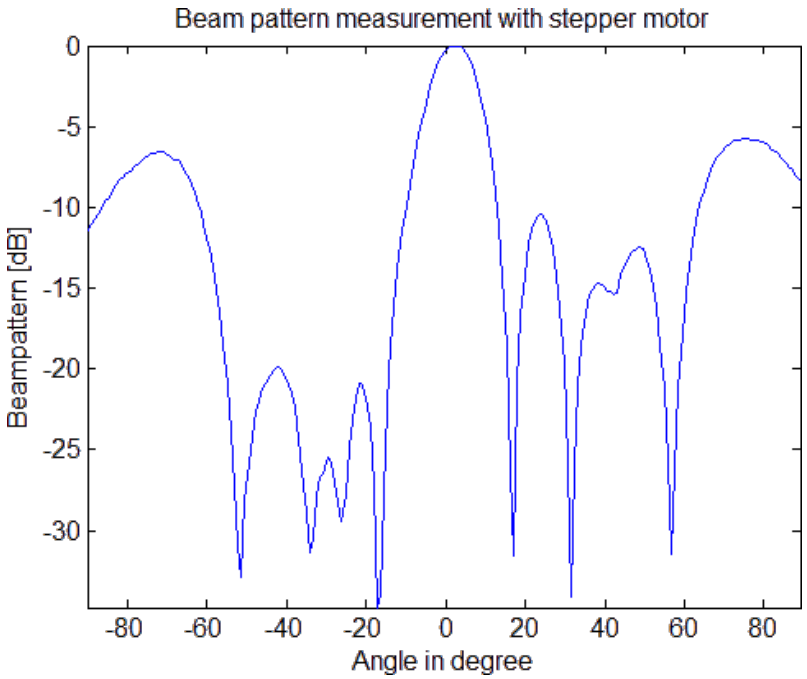

Figure 5.3: Measurement of microphone array steered with stepper motor. Transmitted signal is 8 kHz sine wave.

explained in Section 3.2.1, the plotted phase consists of the relative phase difference due to the position of the array, and a phase error due to microphone unideality. The phase error is present in every position of the array, and zero-angle calibration

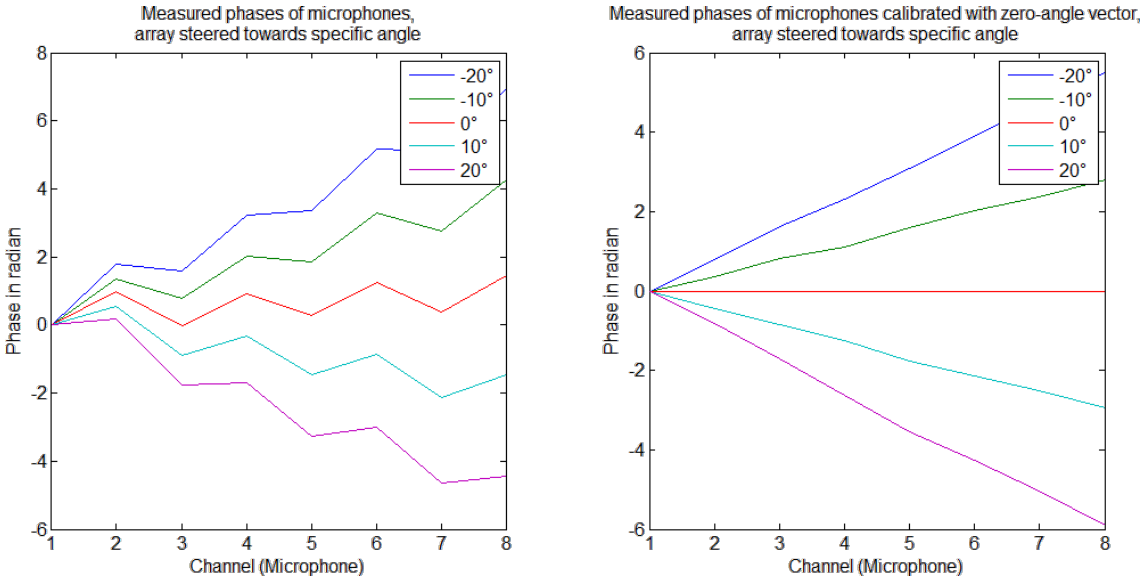

Figure 5.4: Measured phases of the microphones when the array is mechanically steered towards five different angles. Microphone number one is chosen as the reference point and forced to zero artificially. On the left, the microphones are uncalibrated and on the right they have been calibrated with zero-angle vector.

can be used to compensate this. Figure 5.5 presents the phase calibrated version of the same signal as in Figure 5.3. The beam pattern has become closer to the desired shape. The high levels at ±80° have disappeared. However, the side lobe levels remain slightly low and the nulls are barely visible. In order to correct this, the amplitudes should be compensated. The beam pattern, with both the phase and amplitude calibrations, is depicted in Figure 5.6. With the full calibration, the beam pattern begins to resemble the ideal pattern. As the lower side lobe levels mean better performance, the measurement result without amplitude calibration is better than the one with the amplitude calibration. Most likely the error in amplitudes has worked as an appropriate window for side lobe reduction. As this window is not a controlled one, it cannot be trusted that the effect would be always favourable. Thus, the amplitude calibration is also needed.

5.2 Digital Steering

The next measurement phase was to test how the C-2 microphone array would perform in digital steering, or, in other words, as a DS beamformer. A typical motivation for using the digital beam steering is to discover the direction in which the unknown sound source is located, or the direction from which the echo was reflected. In all of the following measurements, the sound source was located in front of the array. This way, the digital steering could easily be verified. Four different tests were conducted in this phase:

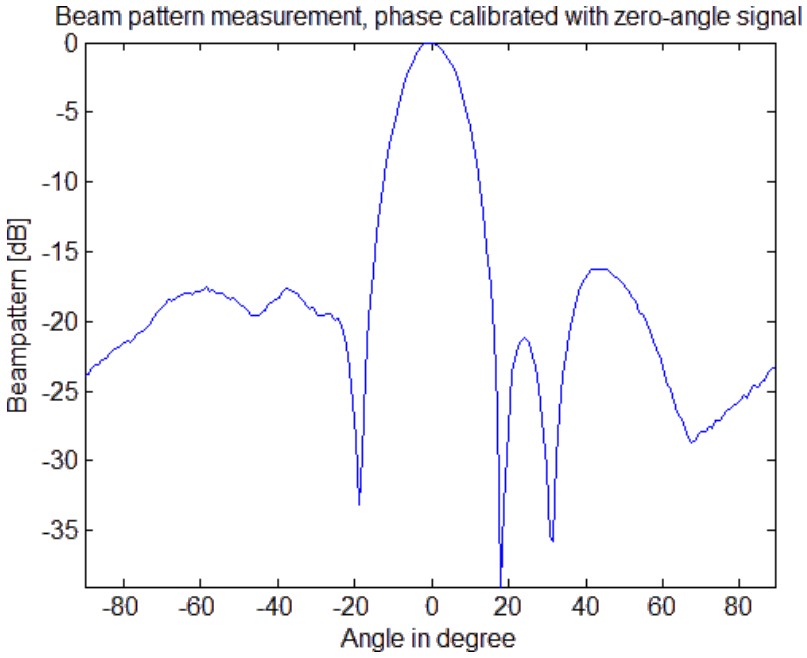

Figure 5.5: Measurement of microphone array steered with stepper motor and phase calibrated. Transmitted signal is 8 kHz sine wave.

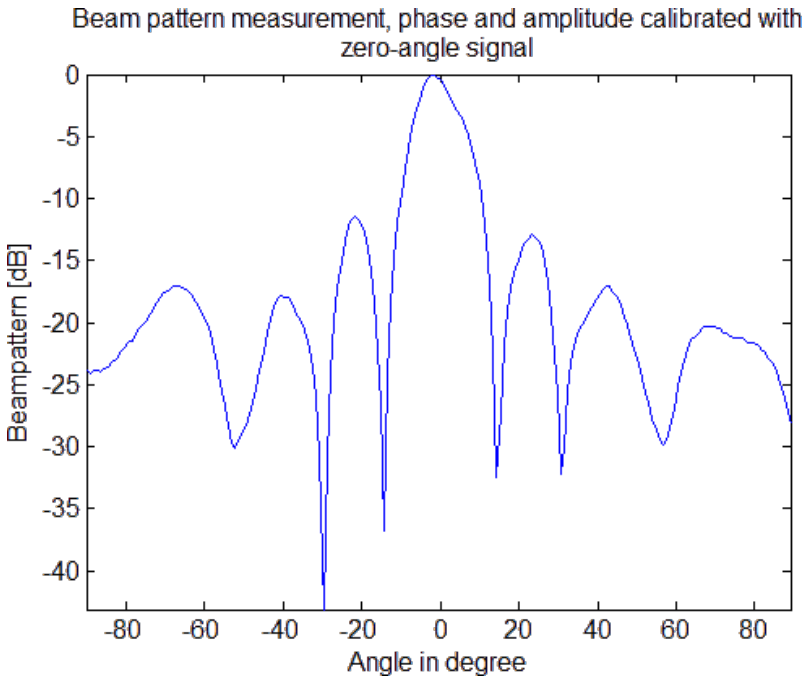

Figure 5.6: Measurement of microphone array steered with stepper motor. Phase and amplitude are calibrated. Transmitted signal is 8 kHz sine wave.

1. Simulations with beam steered towards a different direction than the direction the array was physically pointing towards, while the array was rotated mechanically.
2. Measurements of the previous simulations.
3. The measurement of actual digital scanning with multiple beams pointing towards different directions, while the array was pointing forward the whole time. The calibration values were with 10 degree intervals.
4. The same measurements as the previous test, except that the calibration values between measured values were interpolated.

The first DS beamformer measurement was to place a phase difference to microphones, so that the beam is pointing towards another direction than straight ahead, and then steering the array mechanically. While this is an inpractical way of scanning, it nevertheless is an effortless way to ensure that the microphone's directional pattern is steered to the desired direction. This is depicted in Figure 5.7. The correct phase shift for each microphone was achieved with the calibration measurements. Obtaining the phase shifts was the motivation for different incident angles in calibration measurement. The calibration value for digitally steered beam should correspond to the angle to where it is intended to point at. The theoretical steer to $+20°$ and to $-10°$ is illustrated in Figure 5.8. As was expected, the pointing

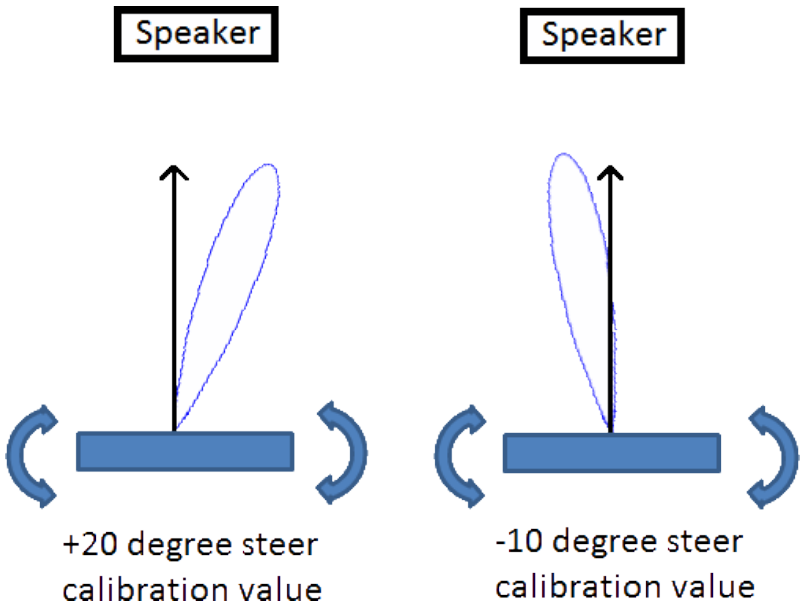

Figure 5.7: Illustration of used setup in Figures 5.8 and 5.9. Beam steered 20 degrees right and 10 degrees left and in the both cases the array is rotated.

direction of the main lobe is not the same as the direction the array is pointing to. Figure 5.9 depicts two cases where the array is calibrated with the calibration measurement values in an angle other than 0°. The beam directions remain the same as in Figure 5.8, and the used calibration values are the ones corresponding to +20° and to -10°. The result was the expected one. The beam patterns resemble the

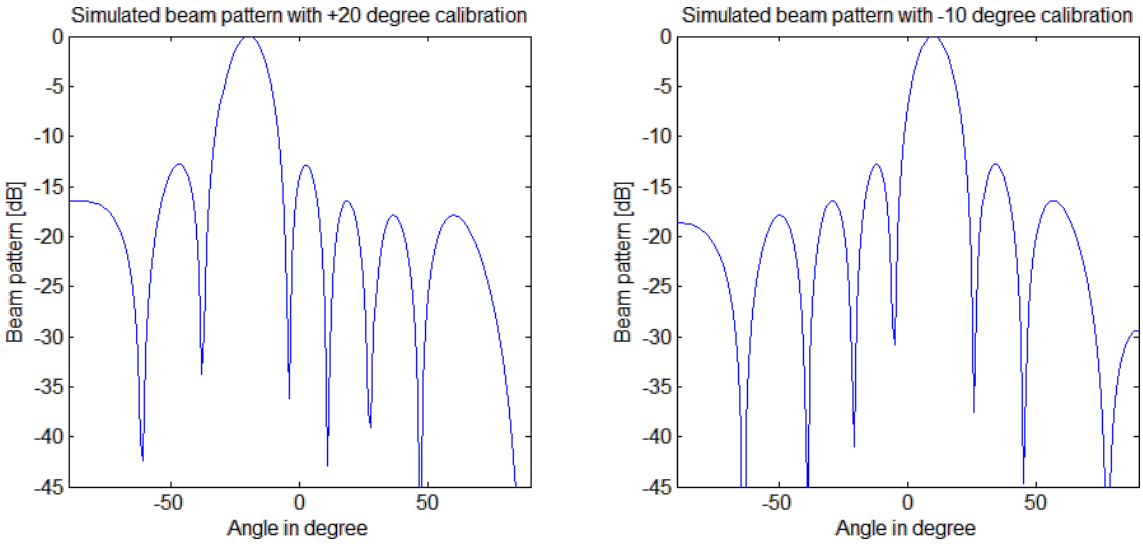

Figure 5.8: Simulations of array steered digitally towards -20° and +10°. X-axis ticks are relative to array's pointing direction and for this reason +20° steer results in beam pattern where the main beam lags the array by -20°. Naturally -10° steer results in beam pattern where the main beam is leading the array by +10°.

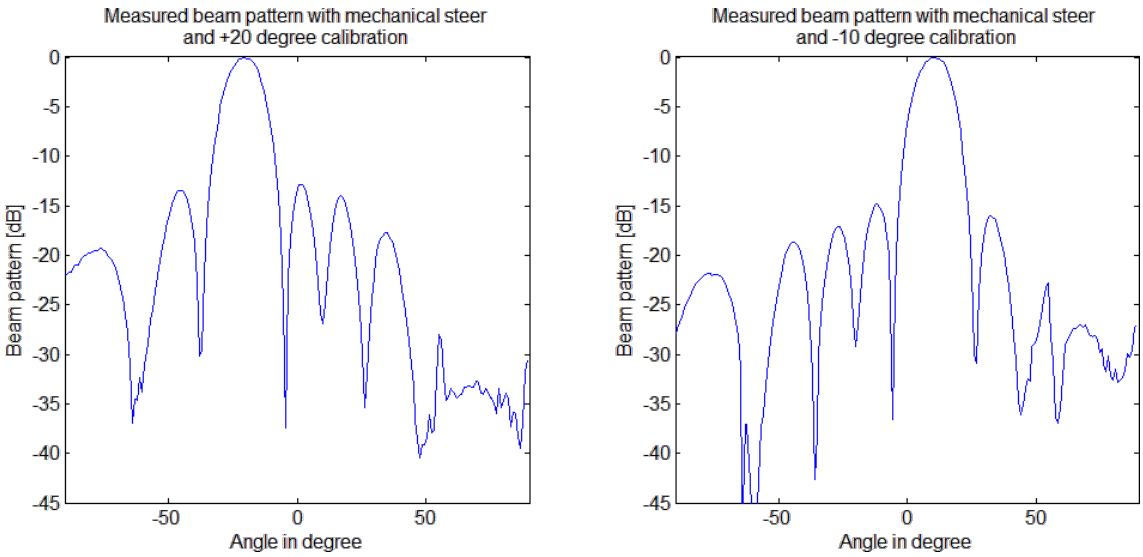

Figure 5.9: Measurements of array calibrated with +20° and -10°calibration values and steered mechanically. Reason for mirrored beam patterns is the same as in the simulations.

theoretical patterns, despite the fact that few unidealities with the sidelobes exist. After verifying the functionality of the used method, the practical digital steering measurement could be conducted.

The actual digital steering is done with several simultaneous beams, which cover the entire area of interest. Thus, the situation is the same as with mechanically rotated array, but the array is steered by using phase shifts. This can be done by measuring all the possible calibration angle values, and then using the value nearest to the angle where the beam is to be formed. This is illustrated in Figure 5.10. The figure does not illustrate any of the actual measurements, as there are only three

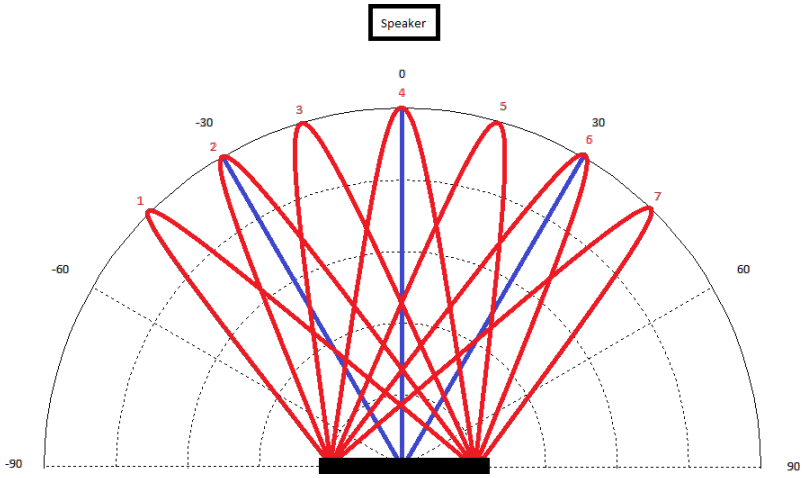

Figure 5.10: Illustration of using sensor array for direction finding. It is done by 7 practically simultaneous beams and 3 calibration angle values.

calibration values and seven beams. Therefore, the precision would be rather poor.

The three calibration values in this illustration are -30°, 0°, 30°: they are marked with blue lines in the figure. Sound source is located at 0° direction which means that it is in front of the array. Seven beams are formed to cover the area of interest and these are marked with the colour red and numbered from 1 to 7. For the beams number 1 and 2 the calibration value used would be -30°. For the beams number 3,4 and 5 it would be 0°. And for the beams number 6 and 7 it would be 30°. For the beams number 3 and 5, values -30° and 30° could also be used. However, normally there is some information about the direction of the echo, and array is already pointing towards that direction. Because of this, it is sensible to use 0° value for these two beams. In Figure 5.11 a measurement similar to Figure 5.10 is presented. It should be noted that when scanning is done this way with the DSP or PC, the scan is so fast that the beams are practically simultaneous. Calibration values were chosen from -60° to 60° with 10° intervals. The sound source was placed in front of the array. On a first glance, the result might seem unsatisfactory. However, the beam pattern has quite clear shape of sinc function. 10° interval would seem to be too long for the application, and the phase calibration is not accurate enough. Yet, even with this configuration the probability of detecting the target at the correct direction is high. A way to improve the calibration without increasing calibration angles would be to use interpolation. The same measurement with interpolated calibration values is presented in Figure 5.12. As expected, the sudden changes in

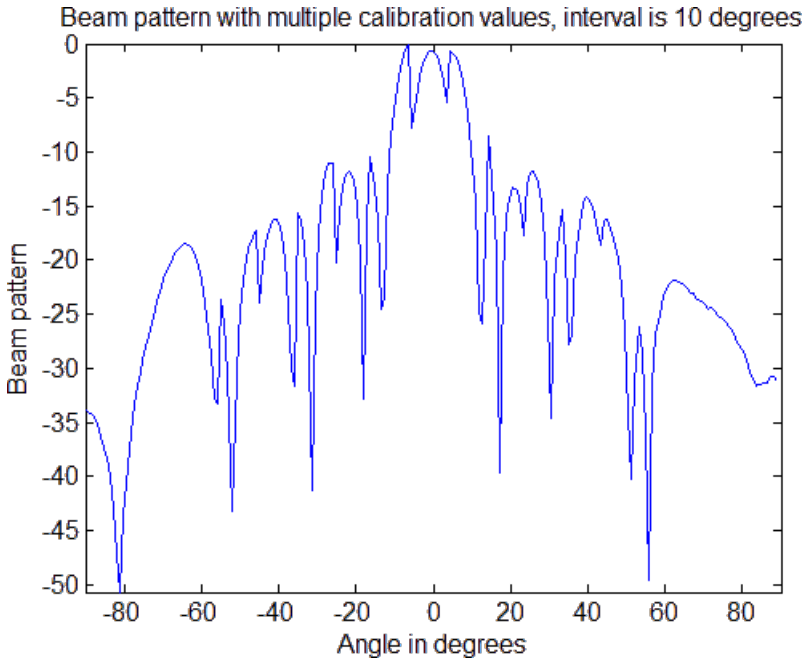

Figure 5.11: Measurements of digitally steered array with calibration interval being 10°. Speaker is located in front of the array.

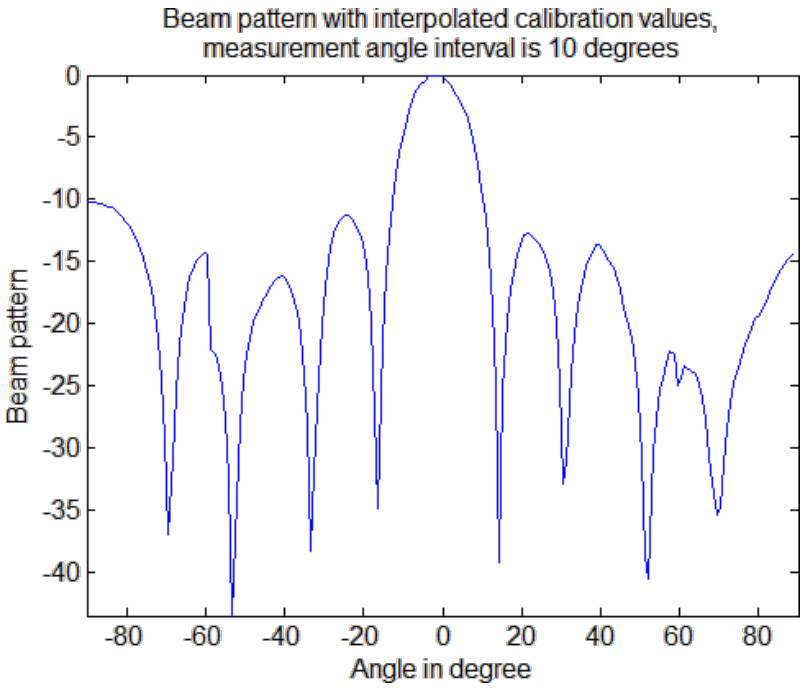

Figure 5.12: Measurements of digitally steered array with interpolated calibration angle values. Speaker is located in front of the array.

the beam pattern disappear. It should be noted that all the calibration values were between -60° and 60 °. This explains the high side lobes outside this band. However, this is not an issue since the practical digital steering is applicable only from -60° to 60°.

5.3 Broadband Beamforming

The last research topic was the use of broad frequency band. A beam pattern of array with 8 microphones was already presented in Figure 3.6. The frequency was from 2000 Hz to 8000 Hz. On the left side of Figure 5.13, a measurement of the microphone array over the same band is presented. On the right side, the same is accomplished with zero-angle calibration. As for the main lobe, even the uncalibrated figure looks reasonable. Side lobes, however, do not have the desired shape. Especially on the higher frequencies, there is a strong lobe level at around ±80°. After the calibration, the sidelobes also take the desired shape, and the overall beam pattern appears the way it should.

As explained in Chapter 3, the frequency dependency of the main lobe is a physical property. The chapter also introduced the concept of nested arrays. The simulations and measurements on the topic of nested arrays will be presented here. The main objective was to produce a constant main lobe for the whole frequency range. The amount of microphones was a severe limitation: thus, only a 2x5 nested

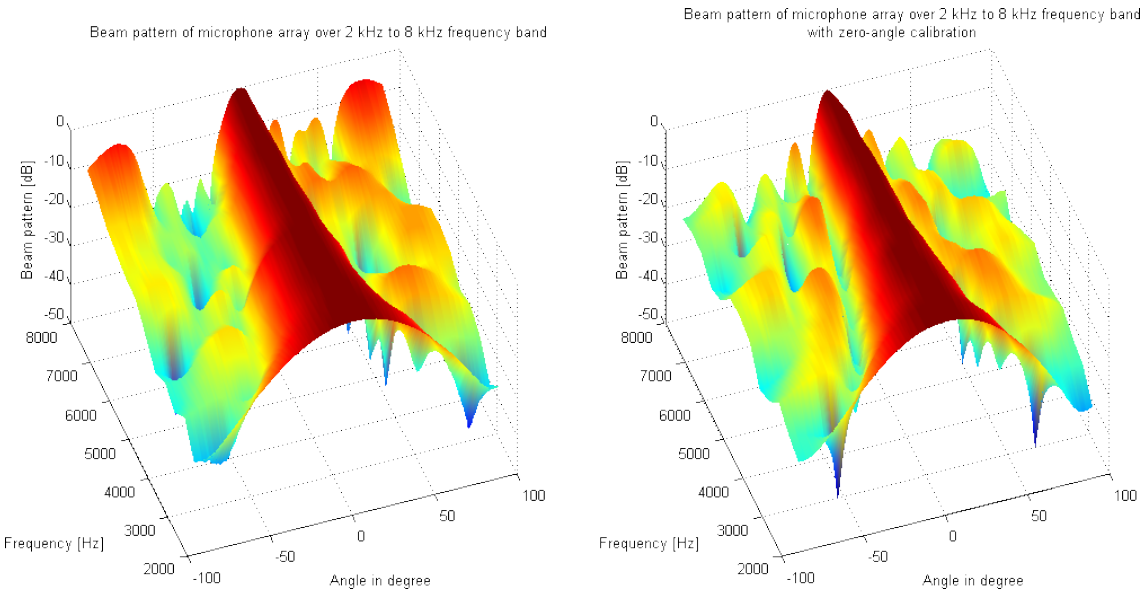

Figure 5.13: Measured beam pattern over frequency range from 2000 Hz to 8000 Hz. Uncalibrated pattern on the left and zero-angle calibrated one on the right.

array could be used. This was, however, enough for verifying the functionality of the nested arrays. In Figure 5.14, a theoretical beam pattern of 2x5 microphone nested array is presented. The spacings for the sub-arrays are 2 cm and 4 cm. The corresponding frequency ranges are from 4 kHz to 8 kHz and from 2 kHz to 4 kHz. With this amount of microphones, the main beam is clearly not even near to

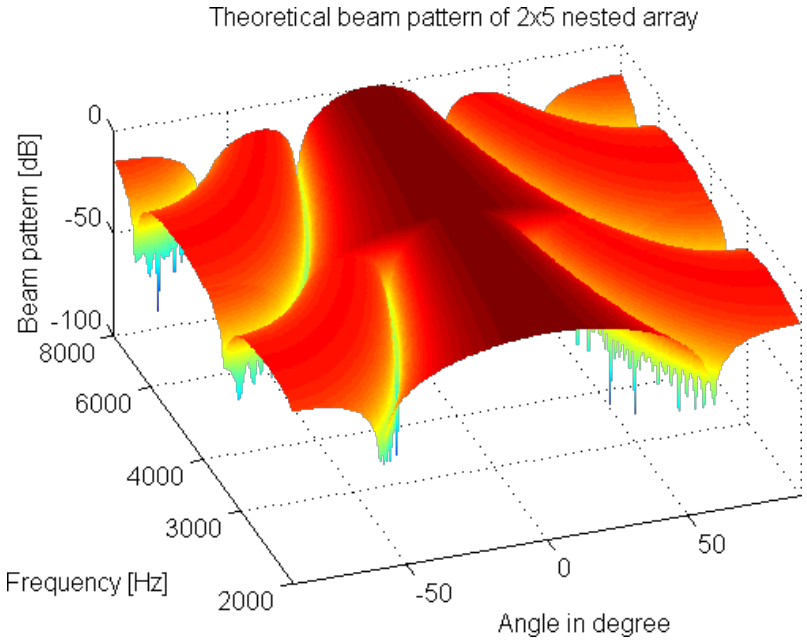

Figure 5.14: Theoretical beam pattern of 2x5 nested array. Spacings are 2 cm and 4 cm. Frequency ranges are from 2 kHz to 4 kHz and from 4 kHz to 8 kHz.

constant. In order to produce an approximately constant main beam, a few filter coefficients were added to the array. This was done so that the coefficients widen the main beam on higher frequencies. The coefficients were decided with brute force technique. First, Dolph-Chebyshev weights were applied. They were then modified manually by trial and error method. The procedure was repeated for every every frequency band, until the beam width became nearly constant. The simulation is depicted in Figure 5.15. As the figure demonstrates, the beam shape is nearly the same for the entire frequency band. With this many microphones the beam is rather wide, but a narrower beam could be achieved with a greater number of microphones.

The final measurements were done to verify the method with actual microphones. In Figure 5.16, a measurement of the 2x5 nested microphone array is depicted. As expected, the shape resembles the theoretical model. The shape of the main lobe is close to theoretical shape, and side lobes are visible as well. Figure 5.17 presents the beam pattern after the same weights that were used in theoretical simulation are applied. As expected, the main beam became wider. Even though minor unidealities are present, the result is nonetheless satisfactory. This clearly verifies that the tapered nested array is a possible solution for producing constant directivity for a sensor array.

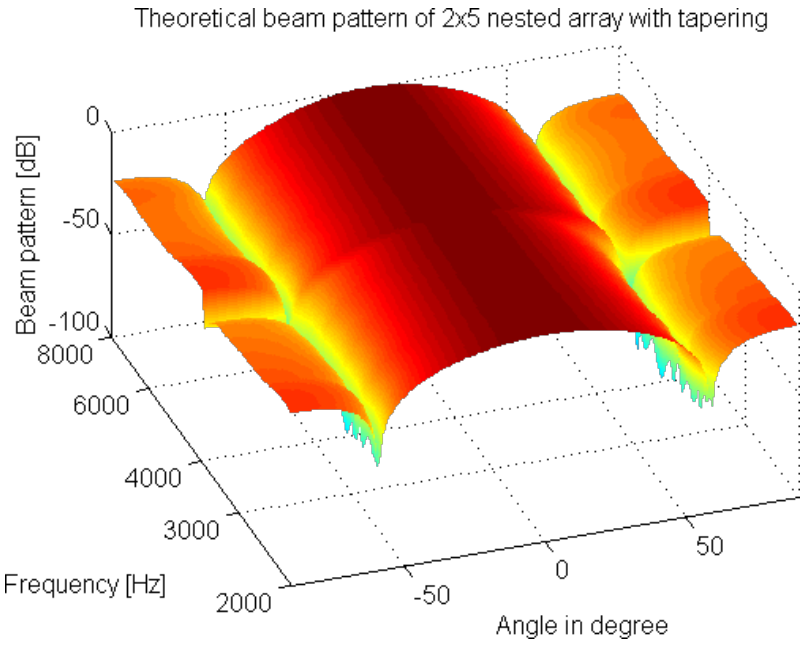

Figure 5.15: Theoretical beam pattern of 2x5 nested array with tapering method. Spacings are 2 cm and 4 cm. Frequency ranges are from 2 kHz to 4 kHz and from 4 kHz to 8 kHz.

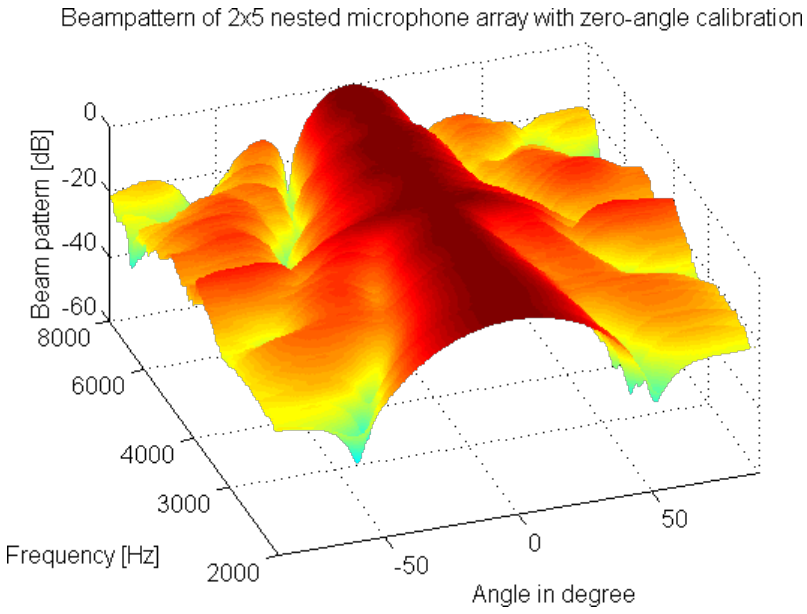

Figure 5.16: Measured beam pattern of 2x5 nested array. Array is calibrated with zero-angle values. Spacings are 2 cm and 4 cm. Frequency ranges are from 2 kHz to 4 kHz and from 4 kHz to 8 kHz.

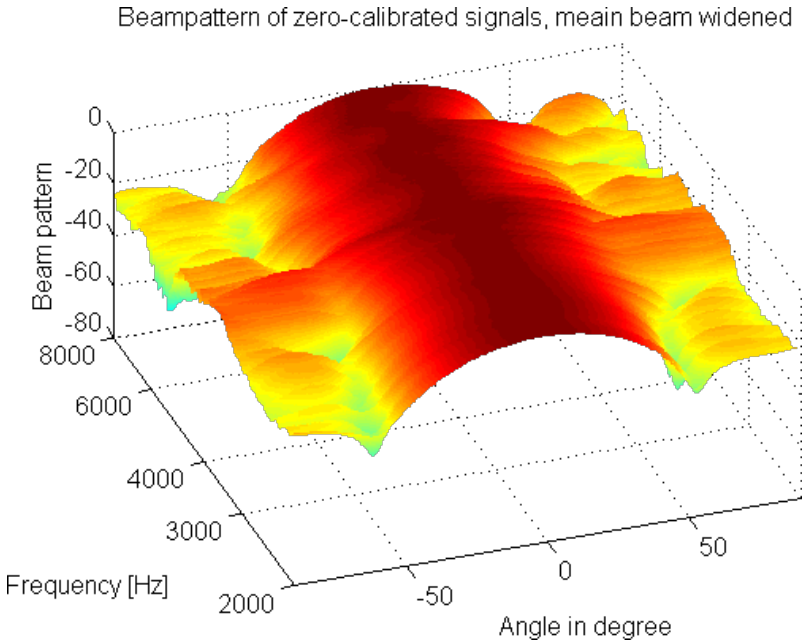

Figure 5.17: Measured beam pattern of 2x5 nested array with tapering method. Array is calibrated with zero-angle values. Spacings are 2 cm and 4 cm. Frequency ranges are from 2 kHz to 4 kHz and from 4 kHz to 8 kHz.

6. CONCLUSION

The main objective of the study was to verify if comsumer microphones could be used in a sodar system, and to gain an understanding of phenomena behind sensor arrays. This was accomplished by simulating the array performance with theoretical sine waves, and measuring the same cases with the microphone array. After the first measurements it became clear that Behringer C-2 microphones could be used in the system. However, practical unidealities caused certain inaccuracies. Thus, phase and amplitude calibrations were deployed. With calibrations the results became satisfactory. The beam patterns of the measured cases were similar to the theoretical models, and veritably proved that Behringer C-2 microphones can be used in sodar. However, the size of the microphones presented certain challenges. The largest element spacing is the width of the microphone, which limits the highest usable frequency. Because the directivity is better on higher frequencies, the diameter of the microphone is the main limitation for the performance of the implemented system.

In this thesis, sodar was approached through a strong analogy to radar. This led thoughts to a DOA estimation: thus, digital beam steering was covered. Again, simulations and measurements were done, and the beam was digitally steered towards the direction of interest.

As the signal frequency is not constant in many of the transmitted waveforms, certain methods for controlling the main lobe beamwidth were necessary. A nested array was proposed as a solution for this. Due to the number of microphones, the resulting main lobe was rather wide. However, with a larger number of microphones the lobe could be narrowed. Nevertheless, it was proved that nested arrays can be used for controlling the frequency dependency of microphone array.

All measurements were conducted in an office environment. This caused problems with echoes for example, and these problems had to be taken into account with the measurements. There were also few drawbacks with parts of the research system. For example, the stepper motor had no feedback which meant that the initial angle for every measurement had to be adjusted manually.

6.1 Lessons Learned

Reflecting back on the project, it is easy to see what could be improved. Possible areas of improvement include equipment, methodology and planning.

The project began with obtaining inexpensive microphones and beginning the measurements. In a normal development process, the initial tests are conducted with better equipment. After the system works with the better equipment, more cost-efficient ways may be devised. Because here initial phase was conducted with inexpensive equipment, all the non-idealities were strongly present in the measurements. Also, the use of anechoic chamber to minimise the effect of environment would have been beneficial for the study.

Generally speaking, the thesis helped the author to gain new knowledge of sound wave ranging and electromagnetic wave ranging. Most importantly, a deep understanding of sensor arrays was obtained as well.

6.2 Future Development

A few ideas concerning the possible uses of the system and ways to improve it are presented below.

An ongoing research on the same sodar system is the use of ISAR algorithm with sound waves. In this application, the waveform used is a frequency modulated pulse: it uses a combination of techniques of CW radar and pulsed radar, which were introduced in Section 2.1. ISAR is an acronym of Inverse Synthetic Aperture Radar: it is radar which makes a two-dimensional image of a target. In other words, this research project aimed to generate a two-dimensional image by using sound waves. This image was successfully created with good results. The objects used in the measurement were clearly recognisable: a precision of less than 5 mm was achieved, making the result more accurate than was expected in advance. At the moment, a conference abstract written by Tomi Nihtilä, Juha Jylhä and Ari Visa on this ISAS research topic is waiting for approval.

Since the size of the microphone is a performance problem, a way to reduce the receiver size should be discovered. Because smaller microphones tend to be more expensive using them is often out of question. Use of on-chip microphones, such as MEMS microphones, could be an interesting research topic. This could also resolve the problem with the number of channels in the audio interface. A data collector would be needed, and implementing the data collector would be a good research topic on itself.

This study proved that Behringer C-2 microphones satisfy the requirements of sodar. However, C-2, a cardioid type microphone, was the only microphone type measured. It would be interesting to measure different kinds of microphones to

investigate which type would be the optimal choice for a microphone array.

The tapering weights of suggested broadband beamforming method were chosen by trial and error method. This is not an efficient method for obtaining the filter weights. Instead, an automated method for finding the weights should be implemented.

All measurements in this thesis were done with the speaker located in front of the array, which equals the operation of a passive radar. It would be interesting to test the basic radar operations of an active radar: to build a speaker array and place it to the same location as microphones. Thus, the system could be used to listen echoes which is a common radar operation.

REFERENCES

[1] Haykin S., Lewis E.O., Raney R.K., Rossiter J.R. 1994. Remote Sensing of Sea Ice and Icebergs. John Wiley & Sons, Inc. 686 pp.

[2] Ainslie M.A. 2010. Principles of Sonar Performance Modeling. UK, Springer-Praxis Publishing. 800 pp.

[3] Van Veen, B.D., Buckley K.M. 1988. Beamforming: A Versatile Approach to Spatial Filtering. IEEE ASSP Magazine. April 1988.

[4] Skolnik M.I. 2001. Introduction to Radar Systems. 3rd edition. New York, McGraw-Hill. 772 pp.

[5] Hansen R.E. 2011. Introduction to sonar. Course material to INF.GEO4310, University of Oslo, Autumn 2011.

[6] Kadygrov E.N. 2006. Operational Aspects of Different Ground-based Remote Sensing Observing Techniques for Vertical Profiling of Temperature, Wind, Humidity and Cloud Structure: a Review. World Meteorological Organization, Instruments and Observing Methods Report No. 89.

[7] About Sodar. Atmospheric Reseach & Technology, LLC. [WWW]. [Cited 16/7/2013] Available at: http://www.sodar.com/about_sodar.htm

[8] Raju G.S.N. 2008. Radar Engineering and Fundamentals of Navigational Aids. New Delhi, I.K. International Publishing House Pvt. Ltd. 443 pp.

[9] Stimson G.W. 1998. Introduction to Airborne Radar. 2nd edition. New Jersey, SciTech Publishing. 576 pp.

[10] Morris G., Harkness L. 1996. Airborne Pulsed Doppler Radar, 2nd edition. Norwood. Artech House, Inc. 510 pp.

[11] Klemola O., Lehto A. 1999. Tutkatekniikka, 2nd edition. Helsinki, Otatieto Oy. 275 pp.

[12] Ghasemi A., Abedi A., Ghasemi F. 2012. Propagation Engineering in Wireless Communication. New York, Springer. 434 pp.

[13] Kraus J.D. 2002. Antennas for All Applications. 3rd edition. Boston, McGraw-Hill. 938 pp.

[14] Rossing T.D., Moore, F.R., Wheeler, P.A. 2002. The Sciense of Sound, 3rd edition. San Francisco, Addison-Wesley. 680 pp.

[15] Karjalainen, M. 2009. Kommunikaatioakustiikka. Helsinki, Akustiikan ja äänenkäsittelytekniikan laboratorio. 240 pp.

[16] Whitaker, J., Benson, B. 2001. Standard Handbook of Audio Engineering, 2nd edition. New York, McGraw-Hill. 872 pp.

[17] Uotila, P. 2009. Tiedonsiirtotekniikan perusteet, kurssimoniste. Tampere. 177pp.

[18] Peltonen, J. 2006. Biofysiikan luentomoniste, luentomoniste. Jyväskylä. Liikuntabiologian laitos. 86 pp.

[19] Mechel, F.P. 2008. Formulas of Acoustics. 2nd edition. Heidelberg. 1213 pp.

[20] Bradley, S. 2008. Atmospheric Acoustic Remote Sensing. CRC Press. 265 pp.

[21] Young H.D, Freedman R.A., Ford L. 2008. Sears and Zemansky's University Physics: With Modern Physics. 12th edition. San Francisco. 1551 pp.

[22] Antoniou I., Jørgensen, E. 2003. On the Theory of SODAR Measurement Techniques. Final reporting on WP1, EU WISE project NNE5-2001-297. Risø National Laboratory, Roskilde, Denmark. 58 pp.

[23] Huttunen, H. 2005. Signaalinkäsittelyn menetelmät, opetusmoniste. Tampere. 115p.

[24] Smith, S.W. 1997. The Scientist and Engineer's Guide to Digital Signal Processing. San Diego, California Technical Publishing. 626 pp.

[25] Levanon N., Mozeson E. 2004. Radar Signals. New Jersey. John Wiley & Sons, Inc. 411 pp.

[26] Valkama M., Renfors M. 2010. TLT-5400 Digitaalinen siirtotekniikka, luentomoniste. Tampere. 452 pp.

[27] Stevenson, A. 2010. Oxford Dictionary of English, 3rd edition. Oxford University Press.

[28] Benesty, J., Sondhi, M.M., Huang, Y. 2008. Handbook of Speech Processing. Berlin. Springer. 1176pp.

[29] Robinson J.E., Matschinski M., Kendall D.G., Chayes F., Hawkings M., Read W.A., Scott E.L., Whitten T.E.H., Graham L. 1970. Spatial Filtering of Geological Data. Revue de l'Institut International de Statistique / Review of the International Statistical Institute Vol. 38. No. 1 (1970). 21-34 pp. International Statistical Institute.

[30] Gonzalez R.C., Woods R.E. 2008. Digital Image Processing, 3rd edition. Prentice Hall. 954 pp.

[31] McCowan, I.A. 2001. Robust Speech Recognition using Microphone Arrays. PhD Thesis, Queensland University of Technology, Australia.

[32] Haupt R.L. 2010. Antenna Arrays: a Computational Approach. New Jersey. John Wiley & Sons, Inc. 527 pp.

[33] Skolnik, M.I. 2008. Radar Handbook. 3rd edition. The McGraw-Hill Companies. 1299 pp.

[34] Haykin, S., Liu, K. 2009. Handbook on Array Processing and Sensor Networks. Wiley-IEEE Press. 889 pp.

[35] Brandstein M., Ward D. 2001. Microphone Arrays. Springer. 398 pp.

[36] Van Trees, H.L. 2002. Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory. New York, John Wiley Sons, Inc. 1443 pp.

[37] Havelock D., Kuwano S., Vorländer M. 2009. Handbook of Signal Processing in Acoustics, Volume 1. New York, Springer. 927 pp.

[38] Benesty J., Chen J., Huang Y. 2008. Microphone Array Signal Processing. Springer. 222 pp.

[39] Cerna M., Farvey A.F. 2000. The Fundamentals of FFT-Based Signal Analysis and Measurement, Application Note 041. National Instruments.

[40] Stranneby D., Walker W. 2004. Digital Signal Processing and Applications, 2nd edition. Elsevier Ltd. 268 pp.

[41] Poularikas A.D. 1998. The Handbook of Formulas and Tables for Signal Processing. Boca Raton, CRC Press LLC. 864 pp.

[42] Eargle J. 2005. The Microphone Book: From Mono to Stereo to Surround - A Guide to Microphone Design and Application. 2nd edition. Oxford, Focal Press. 366 pp.

[43] AudioBox^{TM} 1818VSL Owner's Manual, 2012.

[44] Genelec 1029A Active Monitor System Operating Manual, June 2002.

[45] C-2 Studio Condenser Microphones User's Manual, Version 1.0. 2005.

[46] Nanotec ST5709S1208 Technical Datasheet, January 2007.

[47] Arduino Uno, Arduino. [WWW]. [Cited 17/2/2014] Available at: http://arduino.cc/en/Main/arduinoBoardUno.UwHv5zCXM3w

[48] Arduino Motor Shield, Arduino. [WWW]. [Cited 17/2/2014] Available at: http://arduino.cc/en/Main/ArduinoMotorShieldR3.UwHwIDCXM3w